

An agency that is flexible but flawed

Since its foundation in 1995, the UK Engineering and Physical Sciences Research Council has been radical and innovative in developing its relationships with the communities that it funds. Researchers' distrust of the changes is justified.

THE United Kingdom's physics community is having a rough time. Particle physicists and astrophysicists are feeling the squeeze as the Particle Physics and Astronomy Research Council struggles with the soaring costs to the United Kingdom of international subscriptions. And now the community's other provider, the Engineering and Physical Sciences Council (EPSRC), is turning the screw (see page 755). From a wider historical perspective, the EPSRC's plan may be seen as just one shallow trough on the rippled surface of science funding trends. But it may also be seen as the effect of a flawed experiment in science policy.

The EPSRC is afflicted by conflicts implicit in its mission statement, which requires it to support national wealth creation as well as basic science. At first sight these two objectives should be able to sit side by side, if not comfortably (given funding pressures) then at least compatibly. But such appears not to be the case in practice. Statements by the EPSRC's chief executive make it clear that a relative lack of involvement with industry has counted against the physics community in plans for future funding.

That is symptomatic of a conflict. The EPSRC is the only significant source of grant funding for areas of physics such as condensed matter, atomic physics and nuclear structure. If these areas are not suited to attracting funds from elsewhere, that is all the more reason why the EPSRC should reinforce its commitment to them. But the list of criteria applied in considering programmes is dominated by factors relating to industrial potential, and suggest that if there is no strong industry with which to make a direct link, support for the science will not feed national wealth and is therefore money diverted from a key council mission. That philosophy is a tempting corollary of the UK government's emphasis on wealth creation, but it is dangerous when, as now, it is not balanced by due consideration of science for its own sake.

That lack of balance underlies other signals emerging from the EPSRC, that physicists are not doing as well as others in coming up with good ideas. But on whose evidence is this judgement based? Here one encounters the experimental nature of the EPSRC, which abolished its predecessor's standing disciplinary committees. Grant assessments are now made by anonymous *ad hoc* panels, with minimal continuity, drawn from "colleges" of peers. There is no structure of strategic advisers explicitly drawn from these colleges. The scientific advisers to the council are on the relatively small "technical opportunities panel" which weighs up disciplinary programmes drawn up by respective programme managers — council staff who in turn act as filters of whatever advice and signals they have gleaned from the community.

From the time this system was established, academics have privately expressed a lack of faith in the ability of programme managers, who have been drawn from administrative backgrounds, to represent the long-term interests of their disciplines within such a structure. Certainly, much appears to depend on their presentational abilities. The EPSRC hopes that learned societies will fill some of the (entirely foreseeable) gaps in expert strategic advice, but that has its own problems: what learned society would sensibly start choosing winners and losers from among its membership?

To lose the inertia of standing committees may be an advantage, but not if the flexible and tightly focused structure that replaces it

has critical weaknesses. To enhance trust, EPSRC's senior disciplinary managers should be drawn from their communities — as is the case, for example, in France's CNRS, the US National Science Foundation and the UK Natural Environment Research Council. The council should also be more transparent in its application of troublesome criteria. □

Nature's related journals

The launch of *Nature Biotechnology* makes some clarification appropriate.

THERE are now five publications in the *Nature* family of international journals: *Nature* (founded in 1869), and its monthly relatives: *Nature Genetics* (1992), *Nature Structural Biology* (1994), *Nature Medicine* (1995) and *Nature Biotechnology* — launched as a new title this week.

These journals have several goals in common. The most important is that the papers they publish should be of unusually high quality. It is taken for granted that they should report science that has been carried out to the highest standards. But another aspect of quality relates to unusual impact, and here too all five journals apply exacting criteria. All five achieve these ends by rigorous standards of peer review and rapid publication.

Editorial staff are often asked about the relationship between the journals, and the criteria that govern where a given paper might best be submitted. The following is intended to give some guidance.

As it has always done, *Nature*, appearing weekly, aims to publish papers, from any area of science, with the greatest possible impact, often extending well beyond the discipline concerned. The development of the monthly journals has made no difference to *Nature's* criteria of acceptance in any discipline. The monthly journals aim to publish first-rate papers with exceptional impact within their particular disciplines. More specific details can be found in their guides to authors.

The editors of all five journals make their own independent decisions as to what to publish, and each journal's editor has full responsibility for its content. In turn, the Editor of *Nature* has a responsibility to the publishers to ensure that all publications bearing the word *Nature* in their title maintain the high standards that are expected of them. The majority of papers published in any of the monthly journals are submitted there at the outset. But because of pressure on space, *Nature* has to turn away many papers of very high quality, some of which may well be appropriate for submission to a monthly *Nature* journal. Although the editors of *Nature* may make recommendations to this effect, such suggestions are never discussed with the monthly journals' editors, and it is up to the author whether to act on them. But *Nature's* editors can save significant time by passing on comments of referees to the editor of the journal in question.

More details will be available on the *Nature* web site (<http://www.nature.com>) next month, when the monthlies will also introduce their own pages. □



Physics bears the brunt of research council's accountability demands

London. British physicists are up in arms over a plan by the Engineering and Physical Sciences Research Council (EPSRC), one of the principal sources of support for research in UK universities, to cut their funding two years from now by almost four per cent below its currently projected level.

In contrast, the EPSRC has also provisionally decided to increase funding in the same financial year for both chemistry and mathematics by 5.2 per cent. But marine technology and the 'built environment' have, like physics, both been selected for a cut, in each case by 4.8 per cent.

The proposed redistribution in 1988–89 of the £375 million that the council allocates every year to British researchers — the figures could still be revised before a final decision is reached — represents an explicit attempt to reshuffle funding priorities between different fields of science, taking into account both scientific quality and potential contribution to wealth creation.

EPSRC officials say that the decision to cut the physics budget is intended to send a signal to physicists that, in contrast to chemists and mathematicians, they do not seem to be generating enough innovative ideas, or building enough links with industry, to justify their current funding level.

"This is not a plea to physicists to become more applied," insists Richard Brook, chief executive of the EPSRC, and previously a professor of materials science at the University of Oxford. "We are hoping that physicists will come back with some basic physics proposals that look exciting; if they can do this, then the figures will be changed."

But physicists reject the charge that their field has become static. They claim in reply that the council's move reflects a lack of understanding of the way that physics, in contrast to fields such as chemistry, contributes to wealth creation not through a specific industry but by underpinning developments across a range of industries.

"We are disappointed at the EPSRC's decision," says Alun Jones, chief executive of the Institute of Physics (IOP). He says that the institute is already preparing detailed evidence of the contribution of physics to British industry, and plans to ask its own committees to identify "burgeoning areas" of the subject, to be used to refute the charge that the subject is becoming stale.

The pressures facing British physicists have some of their origins in the government's decision to split the former Science and Engineering Research Council (SERC) in its white paper of May 1993. As a result of

this split, much of SERC's more fundamental research has now been hived off into the new Particle Physics and Astronomy Research Council (PPARC).

EPSRC, which took on most of SERC's remaining activities, retains responsibility for both basic and applied research. But it also has a more explicit mission than PPARC to ensure that the research it funds contributes to national wealth creation (and to enhancing the quality of life).

In line with this new mission, the success of individual scientific disciplines in building links to industry was high among 17 explicit criteria used by EPSRC's new Technical Opportunities Panel — known as TOP, and made up of leading researchers from universities, industry and elsewhere — in preparing recommendations to the full council on how the distribution of research funds in 1998–99 should be adjusted from merely a neutral extension of current funding trends (see next page).

The council, acting largely on the recommendations of TOP, has now suggested increased funding for some areas of obvious application; for example, it proposes that support for research into control and instru-

mentation technologies be increased by 5.2 per cent, and for information technology by 2.8 per cent.

But EPSRC officials point out that — as the proposed cutback in marine technology illustrates — a field's potential contribution to wealth creation is not sufficient to protect its research budget. Conversely, they say, disciplines such as mathematics which have demonstrated both an intellectual vitality and a commitment to closer industrial linkages should be rewarded with the prospect of increased funding.

Thus the council's proposed 3.8 per cent cut in the physics programme, with the implicit threat that the cut could be even deeper if physicists do not pull their socks up, is defended as a deliberate judgement on the current state of those aspects of the discipline — such as solid-state physics and nuclear structure — that have not been taken on by PPARC.

"Many physicists seem to be ignoring the result of the Technology Foresight exercise, and to lack any great interaction with industry," says Brook. He points out, for example, that physicists have done relatively badly compared to those from other ►

Einstein paper on the path to riches

Washington. Who says basic physics does not make money (see above)? A 1912 manuscript by Albert Einstein on relativity theory — the longest and earliest known manuscript on the subject penned by Einstein himself — will be auctioned in New York next month. Sotheby's, which is handling the sale planned for 16 March, expects the 72-page document to fetch between \$4 million and \$6 million, placing it among the most valuable manuscripts ever sold.

Written for an unpublished scientific text commissioned by the German publisher Akademische Verlagsgesellschaft, the document's existence was unknown until 1987, when heirs of the publisher sold it for \$1.2 million to the current owner, an anonymous collector.

Einstein made no effort to preserve his own papers before the 1920s, making this document — written between his development of the special theory of relativity in 1905 and the general theory in 1916 — all the rarer. All but one of the handful of early Einstein papers previously known to exist, the exception being a short paper written

IMAGE
UNAMIGABLE
FOR A COPYRIGHT
FOR REASONS
REASONS

Pride of place: a float representing Albert Einstein in last week's carnival in Rio de Janeiro.

when the scientist was 16 years old, are in public collections in Israel and elsewhere.

Although the edited text of the manuscript being auctioned next month has appeared in print, in a multi-volume set of Einstein papers published by Princeton University last year, the handwritten version contains extensive revisions and deletions that give clues to Einstein's thinking during one of his most creative periods as a scientist, according to David Redden, a senior vice president at Sotheby's. Tony Reichhardt

►disciplines in applications for the so-called ROPAs awards — named after the white paper *Realising our Potential*, and aimed explicitly at rewarding those with a record of successful collaboration with industry.

Some directly involved in the council's decision claim that physicists have failed adequately to 'sell' the value of their work to the new council. "Looking at the general message behind the decision, I feel that physicists need to do more to get across the idea that their work contributes to new developments in many other disciplines," says Laurie Challis, professor of physics at the University of Nottingham, and himself a

17 CRITERIA USED BY EPSRC TO EVALUATE SCIENTIFIC PROGRAMMES

POTENTIAL SOCIO-ECONOMIC BENEFITS

1. Economic Competitiveness
2. Provision of Basic Needs
3. Physical Security and Safety
4. Health Improvement
5. Skills Requirements

NATURE OF RESEARCH

6. Alignment to Foresight
7. Research Potential
8. Pervasiveness
9. Interdisciplinarity

ABILITY TO CAPTURE BENEFITS

10. Strength of User Community
11. Uptake Capacity of User Community
12. Potential for Rapid Technology Transfer

PROVIDER CAPABILITY

13. Strength of Provider Base
14. State of Provider Base

FUNDING CONSIDERATIONS

15. Importance of EPSRC Funding
16. Funding Leverage Potential
17. Capacity to Absorb Increase in Funding

member of the TOPS panel. "This is certainly a glitch, but it may be a useful one."

Others are less charitable, suggesting that the potential value of physics to wealth creation had been misrepresented during the council's deliberations, and criticizing the council for not consulting even those physicists involved in reviewing EPSRC grants applications in reaching its decision. "This was a decision that was reached in some haste behind closed doors," says Jean-Patrick Connerade, professor of physics at Imperial College London.

Jones at the IOP rebuts the implications that physicists are resting on their past achievements, and that an ageing population is reluctant to give way to younger researchers with new ideas. "There is no evidence that physics is anything but exciting," he says, while underlining the institute's keenness to help the EPSRC to identify where physics should grow. "We have said that we will do our best, because if we do not do it, who will?" he says.

But Brook has few apologies for the council's decision. "We have to have an effective mechanism for ensuring that taxpayers get value for the money spent on research," he says. "We are trying to buy the best research we can with the money we have available."

David Dickson

Researchers 'duped' over use of embryos without consent

San Diego. The debate surrounding infertility research at the University of California at San Diego (UCSD) has taken a new twist with evidence that viable human eggs and embryos taken from patients without their consent have been unknowingly used by a research team at the University of Wisconsin.

This was revealed in an audit by the university, completed late last month, of the activities of Ricardo H. Asch, who operated infertility clinics at UCSD and the University of California at Irvine (UCI) before fleeing the country last year while under investigation for alleged misappropriation of embryos and financial offences (see *Nature* 376, 456; 1995).

Asch's research conduct is also under investigation at UCI and Cornell University in Ithaca, New York. Officials from UCI, Cornell and the federal Office for Protection from Research Risks are investigating the transfer of embryos from Irvine to Cornell for experimental chromosome testing.

During the UCSD investigation, auditors from KPMG Peat Marwick found that Asch had in 1994 provided about 450 egg or embryo specimens to a researcher at Wisconsin, Gerald Schatten, without the consent of their donors. Auditors found that at least 21 of the inseminated eggs and 3 embryos were viable. (The remaining eggs and embryos were from failed fertilization procedures, officials said, and would therefore have been discarded as defective.)

Officials at UCSD and Wisconsin say they have no evidence that Schatten was aware that the material sent to him by Asch lacked informed consent. Indeed, records uncovered during the audit indicate that Schatten appears to have been deliberately misled by Asch.

According to Schatten, Asch had provided him with written assurance that the UCSD Human Subjects Committee had approved the use of the eggs and embryos. Schatten has so far declined to comment on the audit's conclusions, citing potential legal action against Wisconsin.

Norman C. Fost, chairman of Wisconsin's Institutional Review Board, says that the university required Schatten to ensure that Asch had obtained informed consent from patients. "We didn't know Asch wasn't telling the truth," says Fost. "We can't hold our researchers to be policemen of researchers at other institutions."

Asch, a native of Argentina, has been living in Mexico City since September after fleeing a federal criminal investigation and at least 18 civil lawsuits. He was not available for comment, and his California attorneys say he is not responsible for the alleged

lapse of research review procedures, blaming university officials for any paperwork mix-ups.

UCSD cancelled Asch's clinical faculty appointment last June, and closed its assisted reproductive technology clinic after the initial allegations. Its subsequent audit has reminded US universities of the sensitivities surrounding research in human reproduction; the National Institutes of Health, for example, have declined to fund human embryo research, because, it is widely believed, of fear of a political backlash.

The UCSD auditors found that Asch had operated his infertility clinic with little administrative supervision, and that there were inadequate written policies to control such operations. Also, the procedures that did exist failed to uncover the alleged irregularities until they were first disclosed by the press last spring.

Both UCSD and Wisconsin officials say that, because of his international reputation as a leading infertility specialist, they relied upon Asch to conduct himself properly. "The affair forces a researcher to question everyone," says one Wisconsin official privately. "This has shaken the entire reproductive medicine field. It will be years until it recovers. It is a shame, because a lot of good science was being done."

Thomas R. Moore, acting chief of reproductive medicine at UCSD, says university authorities "were concerned and dismayed" last year during the early phases of the investigation when they found Asch was using discarded tissue without consent. But these feelings have turned to shock with the recent discoveries. "If indeed this occurred, this is a blatant breach of established policies and procedures as well as an ethical outrage," says Moore.

The UCSD scrutiny of Asch also uncovered another project being conducted without a university-approved protocol. This involved follicular fluid and ovarian tissue from Asch's women patients being given to a university laboratory for unapproved biological and molecular research.

The reproductive medicine researcher conducting these studies, Gregory F. Erickson, was placed on probation for a year, with university officials now monitoring his projects more closely. UCSD officials are also investigating the role of two fellows associated with the project.

"It was an honest mistake to fail to realize I didn't have a protocol in place," says Erickson, who points out his 20-year career at UCSD was previously unmarred. "I am sorry it happened. The university did the right thing by putting me on probation."

Rex Dalton

US and Europe face gap over money for CERN

Washington. Officials from CERN, the European Laboratory for Particle Physics in Geneva, Switzerland, have been visiting Washington this week in an attempt to get negotiations moving on US participation in the proposed Large Hadron Collider (LHC) project.

CERN and the United States agreed last month to set up three working groups to proceed with negotiations which, both promised, would be completed this year. But according to one US official, two of the three working groups are unable to start work because of the gulf between the two sides on the size of the US contribution.

Hubert Curien, president of the CERN Council and a former French research minister, Volker Soergel, Germany's delegate to the council, and Chris Llewellyn Smith, director-general of CERN, were due to have a meeting on Monday (26 February) with Charles Curtis, deputy secretary of the US Department of Energy (DoE), and Jack Gibbons, chief scientific adviser to President Bill Clinton.

Although neither side has said publicly

how much they feel the United States might contribute, the \$400 million recommended by the DoE's High Energy Physics Advisory Panel (see *Nature* 369, 266; 1994) is widely seen as the starting point for the negotiations.

But, as a result of their different accounting methods, \$400 million means quite different things to each side. Indeed, one US official estimates that it would cost the United States \$600–\$800 million to provide what the Europeans would regard as a \$400-million contribution.

CERN expects to spend \$2.5 billion on the construction of the LHC. But that will be in addition to internal spending by CERN and the laboratories of its member countries, whose contribution will come to the project without cost.

To secure the political viability of the project in Europe, CERN would like the United States to contribute as much as possible that can be offset against the \$2.5-billion construction cost. A US contribution, for example, that consisted entirely of steel — with no intellectual content at all — could

be offset to its full value. But a US contribution of software, which might technically be worth more to the project, but which the European laboratories could have written themselves at no cost to CERN, would have no value as an offset.

The width of the gap between the two sides is due to the fact that both would like the United States to make a major contribution not just of hardware but also of expertise. The United States wants an exciting project for its scientists and engineers to work on, and Europe wants as much technical help as possible; but the more expertise the United States provides, the lower the offset value of its contribution.

Llewellyn Smith met US officials in Washington in January (see *Nature* 379, 197; 1996), but Curien missed the meeting because of weather conditions. Officials say that little progress was made in bridging the fundamental gap, but the two sides did agree that detailed negotiations should be conducted by three working groups dealing with administration, the detectors and the accelerator respectively. **Colin Macilwain**

HIV vaccine: sell-out or vote of confidence by Genentech?

Washington. Genentech Inc. of South San Francisco, California, one of the leading US biotechnology companies, after more than a decade of trying to develop an HIV vaccine, has announced that it is setting up a separate company, Genenvax, Inc., to try to bring the vaccine, gp120, to market.

Genentech said last week that it is investing \$2 million to help start the new company, to which it will grant exclusive rights to gp120. Genenvax plans to seek an additional \$18 million in private financing, in order to launch broad-based clinical trials of the vaccine in the United States and Thailand.

But plans to develop gp120 remain controversial. Some critics are concerned that volunteers for large-scale human trials might risk infection with HIV in the mistaken belief that the vaccine is completely effective. Others suggest that the planned tests in Thailand could be open to the charge that a vaccine rejected for phase-3 testing by US officials is being dumped in a developing country.

But Don Francis, an early pioneer of AIDS virology, and president of the new company, is optimistic. "Genenvax is an innovative solution to an important issue," says Francis. "[It] allows for singularity of purpose in developing the gp120 vaccine."

Francis was responsible for Genentech's efforts to develop gp120, which proceeded through phase-2 testing, but were delayed in 1994 when an advisory panel to the National Institute of Allergy and Infectious Diseases (NIAID) recommended against

phase-3 trials (see *Nature* 369, 593; 1994). He disputes the charges made by critics.

"People that call [the proposed Thai tests] a 'dump' are just looking for excuses not to get answers, good or bad, to scientific questions," he says, pointing out that the World Health Organization's Global Programme on AIDS approved efficacy trials in October 1994, soon after NIAID's opposing decision.

Francis also claims that phase-3 testing is now essential for gp120's development. "We are approaching the plateau of our abilities to define potential efficacy in the laboratory," he says. "We've now got to move to empirical real life challenges."

But AIDS activists disagree. "Genentech is cutting its losses," suggests Gregg Gonzales, policy director with Treatment Action Group, an AIDS support group in New York. "Nobody thinks the gp120 products are worthy of going into phase 3 testing."

Industry observers agreed that Genentech is trying to distance itself from a controversial vaccine. "If you're willing to sell something, you're not desperate to keep it yourself," said David Stone, an analyst with Cowen & Co., a New York-based investment bank and brokerage firm. But Francis says that Genentech's decision shows its dedication to gp120. By removing the vaccine from competition with other projects, he says, it has created a company "devoted

IMAGE
UNAMIGABLE
FOR A VARYING
FORECAST
REASONS

Testing grounds: Thailand's red light district could be a prime target for Genenvax's HIV vaccine

and focused" on the vaccine.

Genentech has spent more than \$50 million on gp120 since the mid-1980s. It stockpiled more than 100,000 doses of the vaccine in anticipation of the phase-3 trials that did not materialize. These supplies, along with adjuvants and pre-clinical and clinical data, will be passed to Genenvax.

Genentech will provide \$1 million in seed capital to Genenvax, and an additional \$1 million after \$18 million is raised in private funds. When private financing closes, Genentech will retain a 25 per cent equity investment in the company. Francis says he hopes that private sources will in fact contribute between \$30 million and \$40 million. Genenvax's chairman is Robert Nowinski, a virologist and biotechnology entrepreneur.

Meredith Wadman

Peer review 'still essential', say researchers

Paris. Peer review should be adopted as the principal means of enforcing quality control in electronic science publishing, according to an international meeting of publishers, librarians, scientists and lawyers held in Paris last week. But disagreement remained about the form that such peer review should take in electronic journals.

The meeting, organized by the International Council of Scientific Unions (ISCU) and the United Nations Educational, Scientific, and Cultural Organization (Unesco), agreed that setting up secure, long-lived archives is a priority, and that common standards for indexing papers are needed to ensure accurate retrieval and cross-referencing.

It also concluded that another urgent requirement is for secure digital date stamps to tag original electronic manuscripts, as well as any subsequent modifications made to them. At present, there are no agreed rules for archiving electronic material.

The main disagreement about peer review focused on whether it should continue to be used as the sole criterion for publication of an article, or whether there is also room for electronic archives of non-peer-reviewed material. The latter model has been pioneered by Paul Ginsparg, from the Los Alamos National Laboratory in California, founder of the Los Alamos e-print archives, which are now used by 40,000 physicists worldwide and which carry out 100,000 electronic transactions daily.

Ginsparg himself — who agrees that peer review is generally needed to provide quality control in the electronic literature — supports the concept of a hybrid system, with electronic archives containing submitted papers that have not been peer-reviewed coexisting with electronic journals containing papers that readers will know have been through the review process.

The advantage of such a system, he says, is that scientists would be able to choose between rapid access to non-peer-reviewed information and slower access to that filtered by peer-review, the balance between the two depending on the needs of individual disciplines.

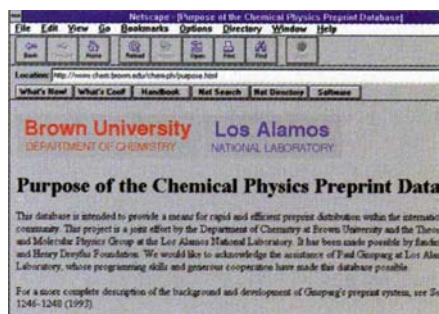
Physicists, for example, have traditionally relied less than other disciplines on peer review, and the Los Alamos archives to a large extent simply replaced the existing tradition among physicists of distributing printed pre-prints. Indeed, Ginsparg says this, combined with the fact that the archives were quickly adopted by an active community, explains the "historical accident" of why the archives escaped the need for peer-review.

The need for electronic publishing to be tailored to the individual needs of particular disciplines was also emphasized by Sydney Hall, a crystallographer at the University of

Western Australia in Perth. Hall says he is "highly suspicious of global solutions" in that they run counter to the flexibility offered by technology.

Scientists at the meeting also criticized the fact that the peer-review process as now practised results in the publication of a single version of particular set of findings. Several scientists argued that, while peer-reviewed papers could be maintained intact with a date stamp, electronic publishing offers the possibility of adding updates or related developments, each with their own time-stamp, on the original manuscript.

The meeting also identified a shift



Electronic revolution: physicists have the tradition of preprints to fall back on.

towards the greater use of raw data, sounds and video in publication, and a lack of common standards for the handling of such material. For example, at least half-a-dozen incompatible protocols are now being used for rotating molecules and other figures.

The meeting stopped short of calling for new common standards to be developed. But it did recommend that "principles and guidelines" be drawn up for what is needed, reflecting the fact that science publishing is only a minor part of the electronic revolution — and that it will inevitably have to adopt standards developed by the wider multimedia industry.

Consensus was reached on many such broad issues. But deep disagreement about the role of publishers, librarians and scientists in electronic publishing remained close to the surface. Many scientists, such as Joshua Lederberg, president emeritus of Rockefeller University, New York, saw electronic publishing, for example, as a means for scientists to regain control over the distribution of scientific information from publishers primarily interested in commercial profits.

Douglas Armati, a UK-based consultant also criticized publishers for holding onto the "dogma" of a "dying transaction paradigm" in their electronic publishing ventures, based on the subscription model that has "served them well and profitably". Indeed, some scientists said they resented the large profits made by many publishing companies from their academic journal

businesses, claiming that such companies often enjoyed monopolies.

Publishers have until now largely defined their operations in terms of the mechanics of distribution and production. But Armati argues that electronic publishing challenges the traditional academic publishing model of obtaining content largely without cost and then selling it back to the contributor base at high prices; publishers will face greater competition and will have to work harder to ensure that customers wish to buy their product.

Indeed, some researchers argued that their main discontent about subscription prices concerns low-circulation, high-cost journals that simply provide papers, and which they felt could be produced more cheaply electronically. High-circulation journals that offer both cheap subscriptions and additional high-quality editorial content have little to fear from the electronic revolution and everything to gain, they predicted.

But many commercial publishers and learned societies at the meeting contested the argument of Ginsparg and others that the new medium will prove much cheaper than traditional publishing operations. Major cost savings would entail researchers taking on much of the work now carried out by publishers, argued David Pullinger, electronic publisher of *Nature*, whereas previous experience has shown that learned societies themselves tend to set up separate publishing operations so that scientists "can get on with research".

Similarly, a representative of John Wiley & Sons Ltd argued that the predicted increase in the electronic archival of raw data would mean that publishers would increasingly have to take responsibility for validating such data, and that this would bring new costs.

Indeed, the general mood of the meeting was that databases validated according to sets of established rules would play an increasing role in scientific communication, with a consequent shift in emphasis from peer review, which is used mainly to assess an author's interpretation of data, to a wider concept of quality control.

Programmes such as the Hubble Space Telescope (HST) and the Human Genome Project will provide useful precedents for validation procedures, argued some, pointing out that astronomers accept, for example, that data from the HST that has been validated by the Space Science Telescope Institute "corresponds to what is happening up there". Others pointed out that such validation procedures will also be important in archiving the vast amounts of declassified information emerging from defence programmes and drug companies which is too copious to be peer-reviewed.

Declan Butler

German regions compete for biotech funds

Munich. Germany last week announced that each of its 17 regions with established biotechnology communities has agreed to take part in a competition to see which of them has the best potential for technology transfer.

The competition is part of a campaign, launched last year by Jürgen Rüttgers, the research minister, to loosen bureaucratic controls on genetic engineering and also to reduce the public opposition that has held Germany far behind its European neighbours, particularly Britain and France.

The rules of the new competition require each region to demonstrate its potential for bringing biotechnological

advances to the market with a minimum of bureaucracy. To be eligible, a region must have an established research base in biotechnology and a sufficient industrial muscle to exploit the results. It must also show that financial institutions, such as banks, are interested in providing funds.

Despite the considerable effort involved in preparing for the competition, the three winners will not be able to look forward to a generous financial prize. They will win only the right to compete with each other for a share of DM150 million in project money that the federal research ministry is hiving off from the DM175 million already available annually for biotechnology pro-

jects. Indeed, there is no guarantee that the three 'strongest' regions will end up with more project money than they would have done without the competition.

But participants in the competition agree that their prime motivation is the prestige that will come from being identified as winners. According to Norbert Koenig, administrative director of the Society of Biotechnological Research in Braunschweig, Lower Saxony, such prestige will help to attract future investment.

A spokesman for the Bavarian economics ministry, which is coordinating Bavaria's application, adds that any region not taking part in the government's competition would be sending the "wrong political signals" to Bonn about its confidence in its biotechnology. In addition, he says, the federal research ministry has remained vague about possible future benefits for winners, and these could turn out to be more significant than the prize now being offered.

Detlev Ganten, director of the Max Delbrück Centre for Molecular Medicine in Berlin, says that it also signals the new recognition among federal politicians of the importance of biotechnology. "In Berlin it has stimulated a lot of discussion between banks, scientists and industry, and the development of many networks, that would not have happened without the competition", he says. In particular, he adds, it has made banks, which have little experience with biotechnology, more positive.

Rolf Kemler, director of the Max Planck Institute for Immunobiology in Freiburg, near the Swiss border, is also enthusiastic about the competition. But he fears that the losers may be seriously disadvantaged, both because there will be less money from the research ministry available to them, and because private finance would tend to concentrate on winning regions, as these would be seen as a lower investment risk.

The competition is being organized in two stages. In the first round, already completed, regions — in some cases as large as a whole *Land*, in others merely a pair of towns — have had to make a general presentation on how they could organize and finance particular research and development ideas.

Regions whose concepts are approved by the research ministry will receive half of the costs, up to a maximum of DM100,000 (US\$68,500) of preparing a full application, due by September. All are expected to receive this grant, and the winners are due to be announced in November. The strongest contenders appear to be Bavaria, Berlin-Brandenburg and Ludwigshafen-Mannheim-Heidelberg, all of which already have strong and well-established biotechnology research bases.

Alison Abbott & Quirin Schiermeir

Fish industry backs seal of approval

London. A leading conservation group has joined forces with one of the world's largest buyers of processed fish to launch an initiative aimed at promoting sustainable fishing. The move is intended by both sides as a bold attempt to employ the power of the market to curb overfishing.

The World Wide Fund for Nature (WWF) and Unilever, the Anglo-Dutch group, have embarked on a two-year consultation exercise aimed at creating a 'sustainably harvested' label for fish. Fisheries organizations, scientists and consumer groups will be encouraged to participate in a proposed independent Marine Stewardship Council, an elected body that will collectively set a standard for the label.

"There is enough scientific information that points to declining fish stocks. The problem is the lack of political will to stop overfishing," says Michael Sutton of WWF.

The initiative, which was launched on the day Unilever announced annual profits £100 million lower than expected, appears to challenge both national governments and the fishing industry. Governments keen to protect employment in the industry have so far failed to prevent overfishing. But if consumers insist on purchasing fish carrying the eco-label, hitherto reluctant suppliers may be forced to comply.

"We are open about the fact that we have a large fishing business," says Caroline Whitfield of Unilever, which has 20 per cent of the frozen fish market in the United Kingdom and the United States. "It is in our interests and those of our employees as well as consumers to make sure that there are enough fish in the sea."

The precise definitions of 'sustainably harvested fish' have yet to be worked out. But they are expected to apply to fish that have not been caught from depleted regions, that are not too young, and that have been caught using equipment that did not trap and kill large amounts of marine

wildlife, known as the 'bycatch'.

All Unilever products — which include Bird's Eye in the United Kingdom and Gorton's in the United States — will do so by 2005, according to Whitfield.

Unilever and WWF will jointly fund the consultation. But the cost of setting up and maintaining the council and of policing the label will be met by conservation groups, governments and foundations, not by industry, says Sutton.

The initiative has so far received a broadly favourable response from the UK Department of Agriculture, Fisheries and Food as well as the UK House of Lords



Select Committee on Science and Technology, which recently published a report calling for the setting up of broad-based fisheries management panels.

But the Earl of Selborne, a member of the committee, warns of potential "cynicism" from the fishing industry.

Trudy Johnston, a spokeswoman for the Scottish Fishermen's Federation, describes the scheme as "interesting", but says "this is the first we've heard of it, and it might have been better if WWF and Unilever had sought direct industry involvement at the outset".

Ehsan Masood

Documents confirm delays in Japanese HIV blood scandal

Tokyo. Japan's new minister of health, Naoto Kan, last week released more evidence confirming his ministry's responsibility in the infection of thousands of Japanese haemophiliacs with HIV through contaminated blood products in the 1980s.

Extracts of documents show that in July 1983, the ministry considered, as an emergency measure, importing from the United States blood coagulants that had been heat-treated to kill viruses. But it quickly rejected the idea, partly to protect the domestic blood product industry, which was behind US and European companies in the development of such products.

Memoranda concerning the proposed emergency imports were found in a 145-page file belonging to Atsuki Gunji, who in 1983 headed the ministry's division in charge of policy on AIDS and blood products. The file and other documents on AIDS were discovered in January in a three-day search ordered by Kan after the ministry had claimed for years, both in court and in the Diet (Japanese parliament), that the documents could not be found.

Much of the information contained in the extracts was reported in a television documentary broadcast by the Japan Broadcasting Corporation (NHK) in 1994 (see *Nature* 367, 584; 1994). But this is the first confirmation by the ministry that emergency imports were considered.

The extracts show that on 4 July 1983 an AIDS study group set up by Gunji discussed importing on an emergency basis heat-treated products from the US company Baxter through its Japanese affiliate Nihon Travonol. According to Gunji's memoranda, the group also considered banning the use of blood plasma imported from the United States — and used in more than 90 per cent of Japanese-made blood products at the time — in the manufacture of non-heat-treated blood coagulants.

But the extracts also show that the ministry was concerned about the possible impact of such action on the domestic blood product industry. A week later, it decided not to make emergency imports, and not to ban the use of US blood plasma. As a result, it was another two years before heat-treated products were approved by the ministry for use in Japan. In the interim, many of Japan's haemophiliacs were infected with HIV.

Gunji, now a professor at Tokyo University, told ministry officials in a written statement last week that he had realized the possible danger of AIDS early in 1983, and that was why he set up the study group. But at the time he thought the infective power of the agent causing AIDS, which had not then been isolated, was low, and that because

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Kan bows to a relative of a patient who died after treatment with HIV-infected blood.

most AIDS cases were in homosexuals, other factors were involved in infection. He therefore felt that the benefits of non-heat-treated coagulants outweighed the risks.

Echoing comments made to *Nature* in 1988, he also pointed out that Japan was being criticized internationally at the time for consuming a large share of the world's blood supplies. In the heat-treatment process, yields of the coagulant are far less than 100 per cent, and Gunji was concerned that use of such products could lead to further criticism of Japan. (In 1988, he also explained his concern that altered proteins in the heat-treated products might set off disastrous antibody reactions in patients.)

Despite these reservations, it was said to have been Gunji who suggested the emergency import of heat-treated products. But, according to the NHK documentary, he gave up the idea in the face of opposition from members of the study group, in particular its head — and Gunji's former professor — Takeshi Abe, who is now vice-president of Teikyo University. In a rare television interview last week, Abe claimed he did not have the power to influence the ministry's decision and was just one member of the study group.

The ministry has also confirmed NHK's 1994 report that in June 1983, just as the study group was being set up, Baxter notified Gunji of the recall of some blood products suspected of being contaminated with blood from an AIDS patient. But, for reasons that have yet to be clearly explained, Gunji did not tell the study group.

The ministry is continuing its internal investigations. Meanwhile, haemophiliacs and their supporters are demanding a thorough independent investigation to establish who decided to drop the idea of emergency imports and why.

David Swinbanks

European deal under fire from US critics over reactor fuel

Washington. A nuclear cooperation agreement between the United States and Europe has come under attack from a broad coalition of US lobby groups as a threat to nuclear non-proliferation.

Paul Leventhal of the Nuclear Control Institute in Washington DC, describes the agreement, which was negotiated last year by the United States and the European Atomic Energy Community (Euratom), as "opening the door to the trafficking of nuclear materials", as it will allow European nations to store and reprocess plutonium and highly enriched materials originating in the United States.

The United States stopped reprocessing plutonium in the mid-1970s, and has since been trying to encourage its allies in Europe and Japan to do the same. In passing the Nuclear Non-Proliferation Act (NNPA) in 1978, Congress sought to impose tougher controls on the allies' use of nuclear materials of US origin. But successive US administrations have regarded good relations with the allies as more important than the strict enforcement of this law.

The agreement with Euratom covers the United States' collaboration with all members of the European Union on civil nuclear research and technology. Although many EU countries perform civil nuclear research of their own, Euratom has responsibility for safeguarding the handling of nuclear materials in all of the countries.

Critics of the agreement hope to attach conditions to its ratification at a hearing of the US Senate's Governmental Affairs committee this week. They will argue that the agreement signed by the Clinton administration breaches the NNPA by failing to retain "consent rights" for the United States over nuclear materials of US origin. The ranking minority member of that committee, Senator John Glenn (Democrat, Ohio), sponsored the NNPA in the 1970s, and is a champion of nuclear non-proliferation.

The critics are also angry with Euratom for negotiating with Russia to buy highly enriched uranium for its research reactors. They say that these negotiations are a direct response to the US embargo on exporting highly enriched uranium, and that the United States should withhold its own agreement with Euratom until the Euratom-Russia deal is blocked.

Groups opposing the agreement include the Nuclear Control Institute, Greenpeace and the Natural Resources Defense Council. They hope to win support from conservatives who feel that the Clinton administration has taken too soft a line with America's European allies in negotiating the agreement.

Colin Macilwain

Engineering president is stripped of duties

Washington. Harold Liebowitz, the embattled and recently elected president of the US National Academy of Engineering, has been relieved of his principal responsibility, acting as vice chair of the National Research Council. The move makes his resignation or forced removal likely in the near future.

In a highly unusual decision, the council of the National Academy of Sciences (NAS) unanimously approved a vote of no confidence in Liebowitz at a meeting two weeks ago, and ordered him to stop conducting business on behalf of the NRC. The NAS charter gives it ultimate legal responsibility for the research council, which is jointly administered by the two academies and the Institute of Medicine.

In a terse statement issued after the meeting, the NAS council said it has noted "with great concern" what it described as Liebowitz's "failure to follow approved procedures of the Research Council, and his adverse impact on Research Council staff productivity and morale".

Liebowitz, who is 71 and a former dean of engineering at George Washington University, took office last July, after narrowly beating a candidate endorsed by the NAE's own nominating committee. He had run for election on an outsider's platform, promising to reform what he viewed as the hide-bound Washington bureaucracy that administers the 32-year-old NAE, which provides advice on engineering matters to the federal government. He also vowed to get more of the academy's 1,800 members involved in its activities.

But Liebowitz's problems since taking office have had less to do with his reformist zeal than with his personal behavior, which sources both inside and outside the academy say has been erratic, confused, and unprofessional. Particularly troubling to the NAS has been his unwillingness or inability to follow rules governing the NRC's dealings with outside organizations, including how it accepts money from government agencies to conduct research studies.

The NAE's own governing council, recognizing problems with their new president almost immediately, curtailed Liebowitz's spending authority and his authority over NAE staff last September, when he had been in office less than three months. But according to NAE and NAS staff members, the aberrant behaviour continued; they are reluctant to give details, partly from fear of possible lawsuits by Liebowitz should he be forced out of office.

After watching the events at its sister organization with mounting concern, the NAS stepped in last week when it concluded Liebowitz had become an embarrassment and a threat to the NRC's integrity. An even terser statement released last week by the Academy of Engineering council stated that

the NAE council "understands the reasons for the NAS Resolution, and its importance to both academies," and was "determining the actions it should take". It promised to inform the members of the NAE of its conclusions "in the near future".

At the beginning of this week, the council's options were still unclear: NAE by-laws contain no specific provision for removing an elected officer, although they appear to allow the council to relieve the president of his duties.

As for Liebowitz, he says he has every intention of serving out his six-year term. In a statement issued on Monday (26 Febru-

ary), he described the NAS resolution as an attempt to "thwart efforts of the NAE president to correct management procedures in the NRC and NAE" and to pursue goals relating to engineering education.

But, having already been relieved of his most important duty, Liebowitz appeared to be in an untenable position. One NAE official said that he expected members of the academy to be asked to vote in a special election for president "sooner or later". The NAE council hopes to meet within the next fortnight to decide their next move in what the official calls a "delicate and very unpleasant situation". **Tony Reichhardt**

Genome laboratory makes its mark

Kazusa, Japan. In a major coup for a young, regional laboratory, the Kazusa DNA Research Institute in Chiba Prefecture east of Tokyo has sequenced the entire genome of a blue-green algae (cyanobacterium) in less than a year. This achievement has surprised many university-based researchers in Japan, whose own sequencing efforts have been hampered by government restrictions on the hiring of technicians, as well by a general lack of funds.

The sequence of the bacterium's entire 3.5-megabase genome will be presented in a paper to be delivered to an international workshop on genomic structure and function in Makuhari in Chiba next month.

Kenichi Matsubara of Osaka University, the leader of Japan's human genome project, says that the group's achievement is very significant for Japanese science. He adds that the results will probably be very significant for plant research as well, given the close relationship between cyanobacteria and plant chloroplasts.

The Kazusa DNA Institute was established by the Chiba prefectural government in October 1994 as the first dedicated genomic research institute in Japan (see *Nature* 372, 3; 1994). Some academics have disparaged the institute and its work as merely a "local effort". But its success in sequencing the cyanobacterium's genome in less than a year has now silenced the critics.

The institute employs 30 scientists and 30 technicians who, in addition to the cyanobacterium work, have also been sequencing large fragments of human cDNA and developing new sequencing technologies. Mitsuru Takanami, director of the institute, says that future plans include sequencing the genome of higher plants.

Michio Oishi, director of the Ministry of International Trade and Industry's (MITI) Institute for Bioscience and

Human Technology, and an adviser to the institute, attributes its success partly to financial flexibility, which has allowed technicians to be hired for the sequencing work, and partly to the focused nature of the institute's projects. Similar sequencing projects at universities are typically spread across several institutions, and staffed by students.

A project to sequence the genome of *E. coli*, which started in Japan in 1989, failed to progress as quickly as the project leaders originally planned because of inadequate funding and also because technicians cannot be hired on university research grants.

Katsumi Isono of Kobe University, who has continued sequencing the *E. coli* genome without direct government funding since 1991, says that repeated appeals to the Ministry of Education, Science, Sport and Culture (Monbusho) to allow hiring of personnel using university grants have been rejected. Very recently, however, Monbusho has indicated that it may establish a more flexible category of grant that will facilitate the hiring of research personnel.

Genomic research in both the public and private sectors has been slow to start in Japan. But recent large allocations from a number of ministries has helped to remedy this situation. In particular, MITI and the Ministry of Health and Welfare (MHW) have separate but related plans to cooperate with the private sector in establishing research institutes dedicated to the development and application of genomic technology and information (see *Nature* 378, 2; 1995).

The site for MITI's planned institute, the Helix Institute, will not be officially announced until next month. But the Kazusa DNA Institute is strongly tipped to house the new institute, and its latest success cannot have done its chances of doing so any harm. **Stephen Barker**

EC group seeks ban on non-medical uses of prenatal screening

Paris. The group that advises the European Commission on ethical issues related to biotechnology last week recommended that the use of prenatal diagnosis (PND) should be restricted to "precise medical indications", and that its use to select sex or other "non-medical characteristics" should be prohibited.

Indeed, the group recommends that such techniques should be restricted to licensed medical centres, irrespective of whether improvements in the simplicity of tests allow them to be offered more widely. The logic of this recommendation is that PND may often reveal information that will require women and couples to take difficult decisions, in particular whether or not to abort a fetus.

In practice, the group recommends that the member states of the European Union should operate PND broadly along the lines on which it is now carried out in the United Kingdom and France, with the emphasis on the licensing of approved centres and on the regular evaluation of the quality of services offered by such centres, in particular in terms of individual genetic counselling and codes of professional practice. □

Ocean programme formally launched

Tokyo. US, Chinese and German scientists signed a memorandum of understanding on Monday (26 February) agreeing to launch a new international programme to drill deep holes in the continental crust. The memorandum was signed at the German embassy in Tokyo by Rolf Emmermann, scientific executive director of the Geoforschungszentrum Potsdam (GFZ) of Germany, Min Zhi, deputy director general of the department of international cooperation of China's Ministry of Geology and Mineral Resources, and Robert Corell, assistant director of the US National Science Foundation. It covers the implementation, management and operation of the Inter-

national Continental Scientific Drilling Program (ICDP).

Scientists from Japan and other countries have also expressed a strong interest in joining, according to ICDP organizers. The aim of the project is to use "the unique capacities of scientific drilling to provide exact, fundamental and globally significant knowledge of the composition, structure and processes of the Earth's crust". ICDP will support drilling costs of selected projects. □

Fusion finds friends in Congress

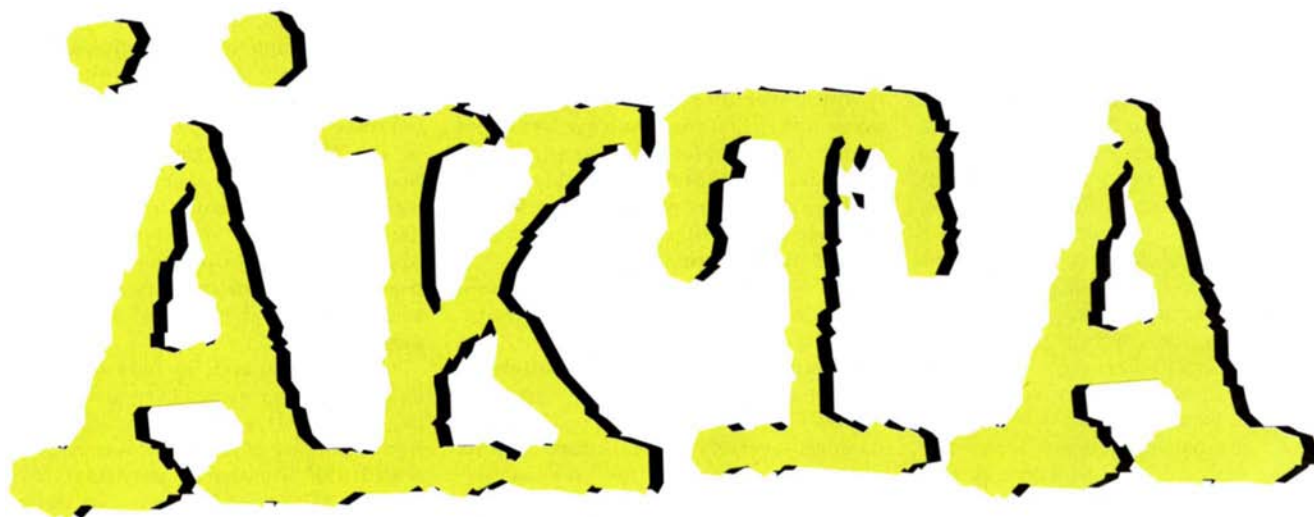
Washington. Fifty congressmen, including some influential Republicans, have joined the Fusion Energy Advisory Committee (FEAC) in calling for funding for fusion research to be raised next year to \$275 million. The Congressmen wrote to Hazel O'Leary, the energy secretary, and Jack Gibbons, President Clinton's science advisor, asking for the money to be included in the 1997 budget, to be published on 18 March.

But the fusion community, which saw its budget crash from \$370 million to \$244 million this year, faces an uphill battle: Robert Walker (Republican, Pennsylvania), chair of the House science committee, thinks it deserves just \$225 million. □

Heaviest element goes nameless

Munich. Scientists at the National Centre for Heavy Ion Research (GSI), in Darmstadt, Germany, announced last week that they had created the heaviest element so far. Element 112 has an atomic mass of 277, contains 165 neutrons and has a half-life of 240 microseconds.

The group will not be suggesting a name for the new element, the sixth they have discovered, until the International Union for Pure and Applied Chemistry (IUPAC) approves their right to name element 108, which they wished to call hassium after the Latin name, Hassia, of the Darmstadt region. A decision from IUPAC about who has the right to name elements is expected in June (see *Nature* **372**, 306; 1994). □



Carbon dioxide emissions

SIR — Wigley and colleagues (*Nature* 379, 240; 1996) believe it would be preferable from an economic viewpoint to defer reducing carbon dioxide emissions for several decades rather than starting immediately. In comparing the economic costs of different reduction schedules, they invoke the cost-benefit criterion of the Framework Convention on Climate Change. They have, however, paid little attention to the 'benefits' aspect, because, they say, estimating the beneficial implications for agriculture, fisheries, biodiversity, human health and so on is a "highly uncertain" exercise.

This is certainly true. Many commentators argue, therefore, that conventional cost-benefit analysis in this complex context is inadequate, as it seeks to assign market valuations to entities that are not normally seen in financial terms, as well as to long-term effects that are often beyond the economist's usual discounting horizon.

Ethicists and political scientists may question both the propriety and political practicality of attempting to pass the cost of reducing emissions to future generations. More fundamentally, the complex, uncertain and potentially unfamiliar consequences of global climate change will require different, more flexible approaches to the assessment of risks and to social decision-making.

Wigley *et al.* ask whether a schedule-related increase of around 0.2 °C in the global mean temperature over, say, a century would translate into "significantly higher damages", and, if so, whether this would offset the economic gains from delay. Most of the Intergovernmental Panel on Climate Change (IPCC) scientists who have struggled, in Working Group II, to identify and quantify the likely impacts of climate change on natural and social systems recognize that such high-resolution arithmetic is impossible. It has been hard enough working at tenfold lower resolution, assessing the likely effects of an overall 2.0 °C rise in global mean temperature over the next century.

In the health impacts chapter of the impending IPCC Second Assessment Report (of which I am the convening lead author), we have identified the major changes in population health that are likely to occur. The effects include deaths from heat waves, altered patterns of vector-borne infectious diseases, some increases in regional malnutrition, problems from water shortages and flooding, and the consequences of population displacement. By their nature, the magnitude and costs of such effects (many of them mediated by climatic disturbances of complex ecological systems) cannot be calculated precisely.

Indeed, thresholds, feedback processes and interactions between the various factors that place stress on the environment could cause surprises.

Attempting to reduce the complexities of climate change, its impact on the biosphere, and the implications for human experiences to market-based values, and then to pass the bill to future generations in the hope that they will appreciate our far-sighted capital accumulation strategy, is not a solution. Cost-benefit analysis is inadequate, and unadorned cost-only analysis more so. The issues posed by global environmental changes will require us to revise our scientific risk assessment and related decision-making.

A. J. McMichael

*Department of Epidemiology
and Population Sciences,
London School of Hygiene
and Tropical Medicine,
Keppel Street, London WC1E 7HT, UK*

Prescribing oral contraceptives

SIR — Michael Rawlins¹ draws attention to errors in your leading article² on the latest UK 'pill scare'. But one error, which he does not mention, acquires fresh relevance with the publication of interim data from the study of cardiovascular disease in oral contraceptive (OC) users directed by Professor Walter Spitzer³.

It is not correct, as your leading article stated, that 'third-generation' OC were introduced to reduce the risk of venous thrombosis; their oestrogen content is essentially unchanged compared to second-generation products, and it is this fact that makes the increased risk of venous disease so surprising. Introduction of third-generation OC, containing desogestrel or gestodene progestagens, was in response to concern about the risks of arterial diseases such as myocardial infarction and stroke. Increased risk of myocardial infarction in OC users was predicted by one of us (V.W.) in 1966 (ref. 4), on the basis of biochemical changes. This possibility was confirmed by a number of epidemiological studies during the 1970s culminating in the Royal College of General Practitioners' Oral Contraception Study, which linked the dose of oral progestagen (norethisterone or levonorgestrel) with arterial disease mortality. At the same time, the progestagen component of some formulations was linked with low serum HDL concentrations. Because high-density lipoprotein (HDL) had been identified as a factor that could protect against coronary heart disease, this

provided a rationale for OC-induced arterial disease. Subsequent reformulations were promoted on the basis of their ability to avoid an HDL-lowering effect. In the latter part of the 1980s the third-generation pills appeared, which increased HDL levels.

Evidence continues to accumulate in favour of a beneficial effect of increased HDL level on arterial disease, and, with the publication of myocardial infarction data from the Transnational Study³, we can now begin to see whether OC-induced changes in HDL may indeed be important. Use of second-generation OC was associated with a significant odds ratio for myocardial infarction of 3.1 (95% confidence interval 1.5–6.3) compared with women not taking OC. The corresponding figures for third-generation products were 1.1 (0.4–3.4). The odds ratio for third- versus second-generation products was 0.4 (0.1–1.2). These data were published prematurely due to the changes in UK prescribing practice resulting from the precipitate action of the Committee on the Safety of Medicines. They nevertheless provide impressive preliminary support for an improvement in arterial disease risk in users of third-generation pills. It is essential to keep these issues in mind in the 'intelligent and well informed collaborative decision making'⁵ that should lie at the heart of modern OC prescribing.

Ian F. Godsland

David Crook

Victor Wynn

*Wynn Division of Metabolic Medicine,
(National Heart and Lung Institute,
Imperial College),
21 Wellington Road,
London NW8 9SQ, UK*

1. Rawlins, M. D. *Nature* 379, 202 (1996).

2. *Nature* 377, 663 (1995).

3. Spitzer, W. *Br. med. J.* 312, 88-90 (1996).

4. Wynn, V., Doar, J. W. H. & Mills, G. L. *Lancet* II, 720-723 (1966).

5. McPherson, K. *Br. med. J.* 312, 68-69 (1996).

Not so public

SIR — "There is no 'burying' of ORI's findings", writes Alan Price of the Office of Research Integrity. "ORI makes its findings very public" (*Nature* 379, 11; 1996). That is not so. Most of the complaints brought to ORI are dismissed without formal investigation (ORI 1994 Annual Report dated April 1995), and these findings are entirely concealed from the public.

For only the few complaints formally investigated does ORI issue public reports, and only if a complaint is upheld does the report identify the people and institution involved.

Charles W. McCutchen

*5213 Acacia Avenue,
Bethesda,
Maryland 20814, USA*

Mt Graham and the environment

SIR—I should like to comment on your report about legislation for Mt Graham's third telescope (*Nature* 379, 199; 1996).

Despite the perceptions of activists to the contrary, the inclusion of language authorizing the US Forest Service to proceed with the third telescope—on the site the service selected as having minimum environmental impact—was in no way at the expense of an HIV/AIDS testing programme for Native Americans. The language mentioned was in the Senate bill but not included in the House of Representatives bill. Before conference, the Senate agreed to accept the position of the House. This all occurred independently of the inclusion of Mt Graham language in the legislation.

Moreover, the Senate language provided no funding for the HIV/AIDS study, but simply expressed the Senate's interest in seeing a report issued. The Indian Health Service (IHS) has full authority to do the study without the Senate language. If telescope opponents are as concerned about this issue as they would have us believe, they would have been working with the IHS all along to address this matter. It appears that telescope opponents are principally interested in creating an issue.

The same is true of other topics. Most of the environmental evidence before the beginning of construction in 1989 and all the evidence since then shows that the observatory has negligible detrimental effect on the endangered red squirrel. This is simply a non-issue invented to try to stop the observatory.

Similarly, the concerns about Native American (Apache) use of the mountain have been propagated by essentially these same opponents of the observatory rather than by the tribe itself. Indeed, the previous tribal chairman felt compelled to write to the Forest Service to explain that these people did not represent the Apaches.

The bottom line remains the same. Opponents of the observatory continue to misuse environmental and cultural laws in order to stop a scientific project that, in itself, will have no significant negative impact on either the environment or on cultural practices.

Peter A. Strittmatter

Steward Observatory,
University of Arizona,
Tucson, Arizona 85721, USA

Using placebos

SIR—I pointed out earlier (*Nature* 375, 530; 1995) that because placebos are unregulated and have not been shown to be inert, an apparent positive, negative or null effect of a drug in a placebo-controlled trial could issue

instead from a negative, positive or same-direction effect of the placebo. J. H. Schoemaker (*Nature* 377, 98; 1995) concedes the point, but believes it prudent to ignore this problem, because facing it might cost money.

Efforts to address the placebo problem need not be costly, and some might even save money. Focus groups could be convened to devise low-cost suggestions. Guidelines on what constitutes acceptable placebos could be developed. Journals could insist that all constituents of placebo and drug be named by weight. One or two placebo agents might be adopted as standards, and their effects examined in detail so that these effects might be adjusted for in analysis. Selected well-studied flavours and coatings could become standard for use with both drug and placebo, actually reducing costs by obviating the need to match the flavour and appearance of the placebo to each drug.

Schoemaker asserts that "in the worst case", if a drug is credited for effects it does not produce because "by chance" there had been negative effects of the placebo, "this may not be harmful". Yet it may be enormously harmful if the result influences clinical practice. In addition, it is naive to suppose that only by chance can a drug be falsely credited because of a deleterious placebo. Pharmaceutical companies currently determine the composition of placebos in trials of their own drugs; as billions of dollars may be at stake in the outcome, production of the 'inert' comparison agent by the drug company that manufactures the treatment drug under study represents a clear conflict of interest, particularly worrying in the absence of regulations, testing of placebo agents for inertness, or even a statement of the placebo composition in published journal articles. These concerns might be mitigated if there were formal regulations about placebo constituents and if production of placebos were removed from the hands of the company manufacturing the study drug.

Beatrice A. Golomb

Department of Medicine,
University of California, Los Angeles,
B-973 Louis Factor Building,
10833 Le Conte Avenue,
Los Angeles, California 90095-1736, USA

Size matters

SIR—Et al. wins Nobel Prize, said Hecht in 1977 (ref. 1). Nothing has changed—multiple co-authorship is the rule rather than the exception. How is the reader to sort out who really did the work, who really had the ideas that lead to a citation classic? Lewison and colleagues propose a fractional grading², which allows a general analysis of a group of papers but does not disclose the true proportional contributions of multiple co-authors to a single paper.

We propose variable font point size as a solution. Briefly, author names will be listed as before, first authorship implying, as now, a major role in the work presented. Senior authors may, as now, run in the anchor position, thus emphasizing administrative success. Both their and other names will be further modified by setting each in a font size appropriate to their contribution. Smith's twofold greater effort will be rewarded by, for instance, 16 point compared to the 8 reaped by Jones. Font sizes

Smith, J. & Jones, S. "Font point size and research contribution" *J. unpubl. Res.* 1, 2-3 (1996).

could vary up to fivefold, easily achieved on modern printers. Thus, scientists would no longer refer to their number of papers but to point size or even to their accumulated point scores, a truer measure of their own work.

It is possible to modify this system, for instance by italics or bold or underlined features. We suggest avoiding such subtleties until the size grading is fully accepted.

This new font point size system would not only help the real contributors but would also be a disincentive to those who were in fact minimal contributors, just added on for the ride—their names in tiny font point letters would show that size really matters.

Eric P. Cohen

Steven E. Mutsaers*

*to whom correspondence should be addressed at:

Department of Medicine,
The Rayne Institute,
University College London,
London WC1E 6JJ, UK

1. Hecht F. *New Engl. J. Med.* 296, 234 (1977).
2. Lewison G., Anderson, J. & Jack, J. *Nature* 377, 671 (1995).

Merely noise?

SIR—The logic of Bill Amos's letter (*Nature* 379, 484; 1996) on 'chance findings' leads to a disturbing conclusion. We know there is a strong bias towards the publication of positive results. We also know that fewer than one in 20 studies are ever published in journals. Could it be that those studies accepted for publication in journals such as your own are merely Amos's chance findings? The logical conclusion is, sir, that your journal may be merely noise ($P < 0.05$).

Christopher Bulstrode

University of Oxford,
Nuffield Department of
Orthopaedic Surgery,
John Radcliffe 11 Hospital,
Oxford OX3 9DU, UK

Correspondence

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to nature@nature.com

Ocean Drilling rules OK?

SIR — An independent review of the Ocean Drilling Program (ODP) that has just been completed contains conclusions that are at variance with those in your recent News article (*Nature* 379, 193–194; 1996). The review was carried out by a committee made up of 11 eminent scientists from seven countries, including Britain and France. It concluded that “the strong record of achievement by the ODP justifies the continuation and enhancement of the programme”, although it also recognized that future success depends on the ability to focus on major scientific objectives.

Throughout its 11-year history, there has been fierce competition from the scientific communities in all member countries, including Britain and France, for drilling time on the drill ship *JOIDES Resolution*.

In the 1995 planning cycle, more than a hundred active proposals at the beginning of the year were narrowed down to a ‘prospectus’ of 17 by August. The December planning committee meeting finally selected only 8: these contain the best science for inclusion in operations in 1997.

Of the 19 member countries participating in ODP, I believe that France is the only one consistently criticizing as “grey literature” the volumes that make up *Proceedings of the Ocean Drilling Program — Initial Reports and Scientific Results*. These represent the basic reference and primary source material for ODP results. The *Scientific Results* volumes are peer reviewed.

These ODP publications are as widely cited as many international journals. Indeed, French scientists are second only to their US colleagues in the number of such citations. A recent review of publication policy by ODP has endorsed these publications, representing the long-term legacy of the programme, as well as providing for more rapid publication of key results in the conventional ‘open literature’ — see, for example, the letter in the same issue of *Nature* (379, 243–246; 1996) by R. J. Behl and J. P. Kennett on their continuing research on core from hole 893 in the Santa Barbara Basin. This important paper relies on the huge compilation of basic data and concepts published in *Proceedings of the Ocean Drilling Program — Scientific Results*, Vol. 146 (part 2).

The Comité Directeur of ODP-France has recommended that France should remain in ODP until the end of the current phase in 1998, and will consider participation in the next phase if certain French concerns are addressed. Many of these changes are already under way.

The proposed Japanese ocean drilling programme, known as ‘OD-21’, is intended to be an evolution and enhancement of the present ODP. The strategy clearly laid out in ODP’s new Long Range Plan involves operation of both a riser drill ship and the

current *JOIDES Resolution* beyond the year 2003.

As for European support for scientific ocean drilling, ODP officials have just completed preliminary discussions with the Secretariat of the European Science Foundation’s European Boards of Marine and Polar Science (see *Geotimes* 41, 9; 1996) on mechanisms for closer cooperation between ODP and future European initiatives in this area.

Finally, I would like to make one correction and one clarification for the record. Russia is not an active member of the programme. The “Consortium” referred to in the article is the Australia/Canada Consortium for Ocean Drilling. On the matter of funding and contributions, the US National Science Foundation pays \$27.7 million of ODP’s total annual operating budget of \$44.4 million, or 62 per cent. Each non-US member contributes \$2.95 million a year towards ODP operations.

David A. Falvey

(Director, Ocean Drilling Program)
Joint Oceanographic Institutions Inc.,
1755 Massachusetts Avenue, NW,
Suite 800,
Washington, DC 20036-2102, USA

The great wave

SIR — The cover of the 18 January issue of *Nature* illustrates the woodblock print “The great wave of Kanagawa”. The attribution is incorrect. The actual artist is Hokusai, not Eijudo. Eijudo was the editor who printed the print. Hokusai was the master who composed this wonderful image, along with approximately 30,000 others throughout his long life. The print was originally drawn in the late 1820s, not in 1831.

Laurence A. Lasky

Genentech Inc.,
460 Pt San Bruno Blvd,
South San Francisco,
California 94080,
USA

On his own head be it

SIR — The breach between religion and science widened yet again with the use of the term “cassock” by D. B. Weishampel (*Nature* 378, 764–765; 1995) to describe the excrescence on the head of *Oviraptor*. It is not a priestly vestment but a helmet that is intended, that is, a casque.

James Murray

Department of Biology,
University of Virginia,
Charlottesville,
Virginia 22903, USA

Reasons to write

SIR — The goal of a paper is to spread ideas and knowledge to a wide group of scientists, which seems so obvious that it should be unnecessary to mention. But I am not convinced that all the authors of the articles in *Nature* and other journals have considered the purpose of their papers.

I think the written communication of ideas in writing has deteriorated lately because of the extensive use of abbreviations and acronyms. Both are used to reduce the length of a text, but at the same time their use leads to a reduction in clarity. For the author and for his close colleagues who use these abbreviations daily, it seems natural and may not cause any problems. But for scientists who hope to broaden their understanding of new ideas in adjacent disciplines, the extensive use of abbreviations can be devastating.

As a referee for various journals, I have noticed that an abbreviation devised for a specific paper is often used only once or a few times. Only a few lines of print have been saved, but the possibility of understanding the paper has been reduced considerably. A limited number of abbreviations for long expressions that recur frequently in a text can reduce the length slightly without jeopardizing comprehension. Authors should, however, be reminded that few readers are probably allowed the luxury of reading a whole paper straight through without frequent and extensive interruptions.

Wibjörn Karlén

Department of Physical Geography,
Stockholm University,
106 91 Stockholm,
Sweden

The great game

SIR — I was interested to read Daedalus’s proposal for the replacement of war by computer simulation (*Nature* 379, 210; 1996) in an issue that also contained a review of Michael Foot’s biography of H. G. Wells (379, 215–216; 1996).

Wells was, of course, one of the founders of modern war-gaming, and in *Little Wars* proposes the construction of “palaces of war”, appointed and upholstered in cork and brimming with lead soldiers, wherein “would-be conquerors and practitioners of *real-politik*” could fight to their hearts’ content, with no little lead widows or orphans to mourn the fallen; “My game is just as good as theirs, and saner by reason of its size,” opined Wells, proposing his “homeopathic remedy for war”.

Rob Paul

NERC Institute of Virology
& Environmental Microbiology,
Mansfield Road,
Oxford OX1 3SR, UK

The inconstancy of the human heart

Christopher Miller

The improper timing of the human heartbeat can now be connected to the arcana of potassium channel gating. What is more, it turns out that a bizarre kind of ion-channel behaviour is in fact quite conformist.

MOST of us take for granted the smooth, regular flow of blood through our veins. But as anyone knows who has experienced an irregular heartbeat (cardiac arrhythmia) or witnessed the catastrophic effects of ventricular fibrillation in a heart attack, the steady lub-dub of the human heart beats perilously close to an abyss of aperiodicity.

The elaborate collection of ion channels in cardiac cell membranes has been designed by evolution to generate a periodic electrical wave and also to preserve its stability in the face of glitches of all kinds—random voltage fluctuations, invasion of electric currents from elsewhere in the cardiac circuitry, and environmental challenges such as drugs and natural toxins. On page 833 of this issue¹, Smith *et al.* describe the molecular mechanism by which one of these cardiac ion channels—HERG by name—acts to protect the heart against inappropriate rhythmicity. In their examination of HERG, they also resolve an irritating headscratcher of a problem concerning the structure and function of these ‘upside-down’ K⁺ channels.

A newcomer to the molecular galaxy of K⁺ channels², HERG was recently identified as the culprit in a genetically determined cardiac arrhythmia³. People lacking a functional HERG gene display an

abnormality on their electrocardiogram called ‘long Q-T syndrome’, which, while causing no distress in itself, predisposes them to sudden cardiac arrest.

HERG’s amino-acid sequence places it squarely in the familiar, extensively studied class of voltage-gated K⁺ (VGK) channels (see box). Typical VGK channels open rapidly (in a few milliseconds) in response to a positive-going (depolarizing) voltage pulse, and then, upon continued application of this voltage, turn off on a slower timescale. This turning-off process, termed inactivation, is observed in all known VGK channels and results from a conformational change that occurs after the channel has been opened by depolarization.

But, surprisingly, HERG does not behave in this way. Very little K⁺ current is elicited by a positive voltage pulse, but if after such a pulse the membrane voltage is brought back to very negative values, large inward K⁺ currents grow over a few milliseconds (Fig. 1). Thus, in comparison to other VGK channels, HERG is functionally upside-down.

A possible explanation for this contrary behaviour⁴ could be that HERG is mechanistically a conventional VGK channel, but that the rates of the activation and inactivation processes are reversed, so that activation is slow (tens of milliseconds), whereas inactivation and recovery from inactivation are, say, tenfold faster.

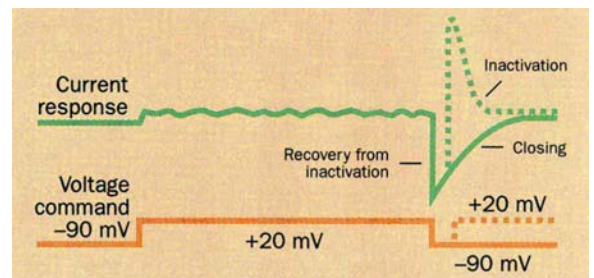
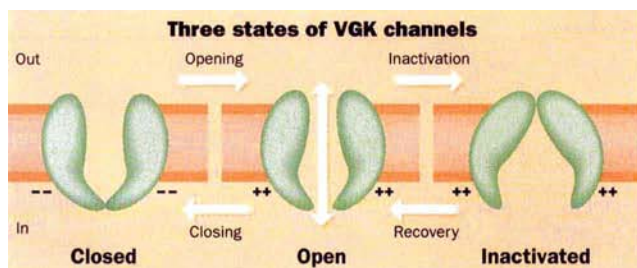


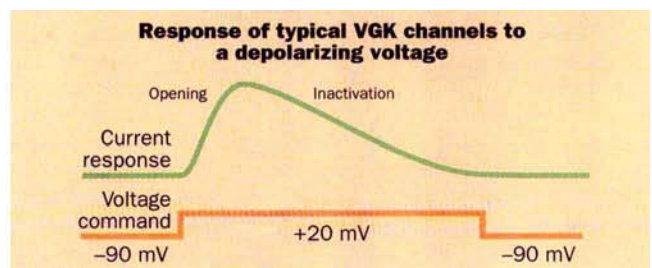
FIG. 1 HERG gating responses. With HERG channels, little K⁺ current flows during a long depolarizing pulse. However, a subsequent negative-going pulse leads to rapid removal of inactivation, and a large inward K⁺ current is seen; if the voltage is quickly returned to a positive value (dashed curves), a large current flows outwards.

A positive voltage would then cause channels to open slowly, but as soon as one opens, it would inactivate. Thus, very few conducting channels would appear at a sustained positive voltage. If the membrane potential were then shifted negative, a fast recovery from inactivation

Voltage-gated K⁺ channels



VOLTAGE-GATED potassium (VGK) channels are found in nerve, muscle and other types of electrically interesting cells. All VGK channels can exist in three basic conformations: closed, open and inactivated, of which only the open conformation can conduct K⁺ ions. (A few VGK channels can also exist in an ‘N-inactivated’ state different from that depicted here.) These conformations are voltage-dependent: negative voltage drives the reactions leftwards towards the closed state, and positive voltage favours the open and inactivated states. Conventional VGK channels are often studied by holding the



membrane at a highly negative voltage, where all channels are closed, and then applying a positive-going, or depolarizing voltage pulse. As channels open over several milliseconds in response to the test pulse, outward K⁺ current rises, but then after a longer time (tens to thousands of milliseconds, depending on the particular channel), inactivation sets in, and the K⁺ currents fall. The detailed shape of the K⁺ current time course reflects the balance between the forward rates, opening and inactivation, and the backward rates, closing and recovery from inactivation.

C. M.

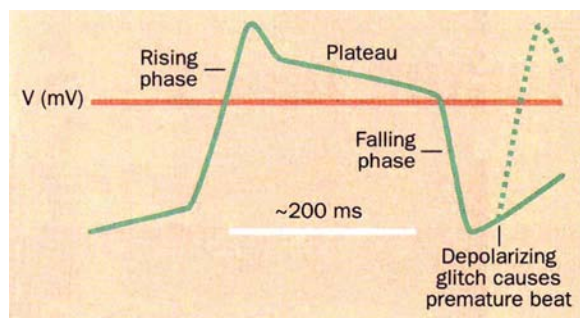


FIG. 2 A normal cardiac electrical waveform (solid green line) and the generation of a premature afterbeat (dashed line) by a depolarizing glitch.

should take place, apparently making the channels open rapidly at negative voltage. (In fact, they were already 'open' before the negative voltage pulse, but were rendered non-conducting by inactivation.)

The new results¹ confirm this idea in three ways. First, the authors contrive to catch the open channels at positive voltage. After a long depolarized interval like the one described above, they briefly switch the voltage to a negative value — just long enough to remove inactivation. Then, they re-apply a positive voltage and — *voilà!* — observe familiar, VGK-like outward current through the HERG channels, caught *in flagrante* for a few milliseconds before inactivation sets in (dashed trace in Fig. 1). With this assay, the voltage-dependent activation curve of HERG can be measured, untainted by inactivation, and comes out as expected for a conventional VGK channel.

A second test exploits the pore-blocking action of the cation tetraethylammonium, which in standard-issue VGK channels antagonizes inactivation⁵. With this blocker present, HERG inactivation becomes slower than activation, and now the channel behaves like a conventional VGK channel. Finally, the authors construct an HERG variant carrying a mutation that appears to eliminate inactivation: once again, the channel reverts to normal VGK voltage-dependence.

These careful mechanistic studies nicely illustrate the point that, with just a few adjustable parameters, you (or evolution) can make an elephant wag its tail; the gross behaviour of HERG is upside-down not because of any fundamental difference from other VGK channels, but merely by virtue of a few unusual rate constants in a wholly conventional gating scheme.

Do these minutiae of channel gating have any significance for the role of HERG in cardiac function? That is a difficult question to answer definitively, thanks to the complicated choreography of the many ion channels participating in the heart's electrical dance (performed in Fig. 2). In a normal cardiac waveform, the membrane voltage swings positive in response to the inward flow of Na^+ and Ca^{2+} ions through their respective chan-

nel proteins. This voltage change causes K^+ -specific channels to open after a delay, and the outward flow of K^+ ions returns the membrane to a negative voltage. A healthy heart-beat demands a delicate balance of inward Ca^{2+} and Na^+ and outward K^+ currents to fine-tune the precise shape of the waveform. Sometimes, glitches occur that tend to set off the next beat prematurely, as in certain arrhythmias. The

behaviour of HERG described here, however, is just what is needed to guard against such early 'after-depolarizations'.

A depolarizing glitch just following the falling phase of the cardiac action potential resembles the voltage-clamp protocol used by Smith *et al.* to unmask HERG at depolarized voltages: a long depolarization (plateau phase), followed by a brief repolarization (falling phase), followed by a second depolarization (the glitch). Under

these conditions, HERG channels, happily asleep at other times during the waveform, would quickly awaken to provide outward current and suppress the potentially catastrophic effects of the glitch.

It may be for this reason that patients with HERG channels that have been knocked out of action, either by genetic disruption or, ironically, as a side-effect of certain anti-arrhythmic drugs⁴, become susceptible to sudden cardiac failure. Whatever the particular role of HERG in cardiac pathology turns out to be, K^+ channels can now claim an inherited molecular disease all of their own. □

Christopher Miller is in the Graduate Department of Biochemistry, Howard Hughes Medical Institute, Brandeis University, Waltham, Massachusetts 02254, USA.

1. Smith, P. L., Baukrowitz, T. & Yellen, G. *Nature* **379**, 833–836 (1996).
2. Warmke, J. W. & Ganetzky, B. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3438–3442 (1994).
3. Curran, M. E. *et al. Cell* **80**, 795–804 (1995).
4. Sanguinetti, M. C. *et al. Cell* **81**, 299–307 (1995).
5. Choi, K., Aldrich, R. W. & Yellen, G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5092–5095 (1991).

EARTHQUAKES

Being struck by a balance

Ross S. Stein

HIKING down windswept Solitario Canyon, ground zero of the proposed US nuclear waste repository in Nevada, James N. Brune was struck by the beauty of a huge, precariously balanced boulder. "But why hasn't an earthquake knocked that thing off?" he thought. Or could the rock have been poised to fall for only a short time? Brune got his answer a few days later at the annual meeting of the Geological Society of America, where evidence was presented that the weathered pavement of Yucca Mountain had survived for 30,000 years, long enough to have experienced whatever shaking the local earthquakes had to offer¹.

Brune soon found that such 'precarious rocks' could be dislodged by ground shaking with just 10 to 20 per cent of the peak motions recorded during the recent Landers, Northridge, or Kobe earthquakes. With these insights, Brune transformed a stone of beauty into a stone of witness, extending one aspect of the record of earthquake shaking in the United States by a hundredfold over our brief history, and in the process challenging theoretical ground-motion predictions

in closely watched southern California².

Precarious rocks form underground when joints — or sets of cracks — in granite or welded volcanic rocks (or 'tuffs') slowly widen, isolating the interior core-stones. Erosion eventually washes away all but the core-stones, and occasionally one stone remains piggybacked on another. Stones smaller than about 0.3 m get knocked down by the wind, leaving the large ones to await nearby earthquakes or errant teenagers for the *coup de grâce*. Of the 18 fields of precarious boulders in



Does the boulder in the foreground ensure the safety of the buildings in the background? Just 40 km from the southern San Andreas fault, near Victorville, California, a field of precariously balanced rocks testifies to 10,000 years of seismic sleep.

James N. Brune

California and Nevada that Brune has found by cruising the highways with binoculars in hand, none is in a site that has sustained ground accelerations of more than about 0.2g from historical earthquakes or underground nuclear blasts.

Brune and his students developed techniques to estimate the acceleration needed to topple precarious rocks by subjecting styrofoam rock models to recorded earthquake accelerations, and by dynamic two-dimensional computer modelling³. They also felled rocks in the field by static loading (a geologist's version of killing animals in the service of human medicine?). They found that precarious rocks fall when horizontal acceleration exceeds 0.1–0.2g, in accord with studies of earthquake-induced rockfalls⁴. Semi-precarious boulder fields, with less spectacularly poised boulders, can sustain motions of 0.3–0.4g (ref. 3). They also found a dependence of rocking frequency on the square root of height, with dominant frequencies in the range 1–5 Hz. This is the range for one- to twelve-storey buildings, so serendipitously the boulders stand in quite well for human edifices threatened by shocks.

With help from other geologists, Brune found that most rocks have held their balance for at least 10,000 years, ensuring that they are unlikely to have escaped the largest events that occur at a site. The trick to divining the antiquity of the rocks is to measure not the time a boulder has been on the ground, but the time it has been poised to fall. For that, one must work at the fulcrum. Most boulders are coated by rock 'varnish' layers of manganese and iron oxides. Existing radiocarbon dating and new techniques that associate the varnish layers with climatic events yield a range of 10,500–22,000 years for the periods over which different rocks have been perched.

Because granite and volcanic tuffs only cover a minority of the surface in the western United States, and boulders disintegrate rapidly in all but arid climates, it is difficult to make a systematic survey of precarious rock sites. Brune and his entourage have yet to canvass park rangers, drive country backroads, or take to the air in ultralight aircraft to search for new sites. But even with these limitations, the impact of the rocks promises to be heavy.

If the 1995 working group report on earthquake probabilities⁵ is correct, all boulder fields in southern California should have been flattened during the past 3,300 years of shaking. Instead, most fields have survived at least three times

as long. The report, a seminal document for the incorporation of palaeoseismic, seismic and geodetic data, postulates random, large, rare earthquakes throughout the region to reconcile earthquake occurrence with plate-tectonic motion, which requires a rate of large earthquakes (of magnitude 6 or above) twice that observed during our scant 150-year record. Between the major faults, the accelerations predicted by this model are nearly twice as high as the limits implied by precarious rocks. If instead the 'wildcard' earthquakes are confined to the major faults, quiet zones emerge in between to

accommodate Brune's rock sentinels.

Because our historical record is far too brief to reflect the richness of earthquake occurrence, a 10-millennium record of shaking, no matter how faintly etched on the landscape, is invaluable. "Why didn't I think of that?" is the question that comes to one's lips. Brune not only thought of it, but pursued its many consequences until he could use the rocks to test, if not topple, our best models. □

Ross S. Stein is in the United States Geological Survey, 345 Middlefield Road, Menlo Park, California 94025, USA.

GENETIC CODE

A case for *trans* translation

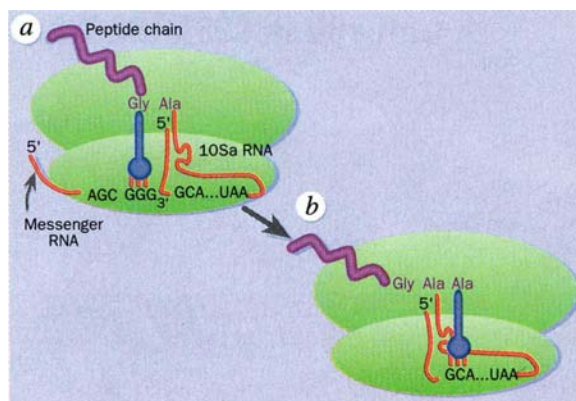
John F. Atkins and Raymond F. Gesteland

THE discovery of single protein monomers translated from two separate RNAs, announced in papers in *Journal of Biological Chemistry*¹ and *Science*², has startling implications for our view of how the genetic material is decoded into a protein sequence. Typically the ribosome machinery translates an RNA sequence from a start site to a stop codon on one molecule of messenger RNA, so that the protein product reflects a strict, linear read-out according to the rules of the genetic code. The realization that the ribosome can switch from one RNA molecule to another comes from the finding of a common terminal 11-amino-acid tag at the carboxy terminus of a number of proteins that are destined for degradation in *Escherichia coli*. The last 10 amino acids of the tag are decoded from the 363-nucleotide RNA, 10Sa, a stable non-ribosomal RNA. The first of the 11, the 'junction' amino acid, has no apparent codon.

These reports from the laboratories of R. J. Simpson¹ and R. T. Sauer², along with work from a group led by A. Muto (H. Himeno *et al.*, personal communication), point to an intriguing mechanism whereby the 10Sa RNA acts both as transfer RNA, which is responsible for transferring amino acids to the growing polypeptide chain, and as messenger RNA. The same 11-amino-acid carboxy-terminal tag was found on truncated murine interleukin-6 synthesized in *E. coli*¹, on lambda cI and on cytochrome *b*-562, when translated from mRNAs without stop codons², and

on polyphenylalanine made *in vitro* with polyuridylic acid as a template (H. Himeno *et al.*). Genetic experiments show that the 10Sa RNA is required for tag addition but, surprisingly, not just for the encoded peptide sequence^{1,2}.

The 10Sa RNA is unusual in that the 3'



A model of *trans* translation based on Keiler *et al.*², in which the 10Sa RNA acts both as messenger and transfer RNA. *a*, 10Sa RNA in the A-site positions alanine for peptidyl transfer. *b*, Decoding switches to 10Sa RNA. Note that the alanine on 10Sa RNA is incorporated without recognition of a codon. If codon-anticodon pairing in the A-site were crucial for transpeptidation, an 'anticodon'-like sequence elsewhere in the 10Sa molecule could pair with the GCA alanine codon. This pairing would need to be disrupted to allow the incoming alanine tRNA to pair with GCA. 10Sa RNA is also involved in modulating the activity of DNA-binding proteins⁷.

and 5' ends pair to form a partial tRNA-like structure which is charged with the amino acid alanine^{3,4}. The alanine from this 'tRNA' is inserted as the first amino acid of the tag without involving the usual steps of decoding. The model proposed by Sauer's group (see figure) suggests that ribosomes poised at the 3' end of an mRNA without a stop codon bind 10Sa RNA charged with alanine (at the A-site of the ribosome), then this amino acid is donated to the stalled peptide chain. The 10Sa RNA somehow folds to fit its internal sequence into the

1. Whitney, J. W. & Harrington, C. D. *Geol. Soc. Am. Bull.* **105**, 1008–1018 (1993).

2. Brune, J. N. *Bull. seism. Soc. Am.* (in the press).

3. Shi, B., Anoshehpour, A., Zeng, Y. & Brune, J. N. *Bull. seism. Soc. Am.* (submitted).

4. Keefer, D. K. *Geol. Soc. Am. Bull.* **95**, 406–421 (1984).

5. Working Group on California Earthquake Probabilities, *Bull. seism. Soc. Am.* **85**, 379–439 (1995).

mRNA track, facilitating cotranslational switching of the ribosome to a new 'mRNA'. The last 10 amino acids are then translated from 10Sa RNA to a normal terminator.

The tag specifies the chimaeric protein as abnormal, targeting it for degradation². The mechanism contrasts with that for regulating protein turnover in mammalian cells, where post-translational sequence-dependent tagging with ubiquitin is followed by degradation in the proteasome⁵.

Elucidation of the details of how a ribosome transfers from one mRNA to another will be fascinating. The 10Sa RNA holds the key, as it must fit into the ribosomal A-site in a special way that allows transfer of the alanine. It then becomes an mRNA, rescuing the stalled ribosome. (Perhaps it should be renamed 'tmRNA' to reflect both these functions.) This process provides a mechanism to recycle ribosomes that have translated aberrant mRNAs, truncated before stop codons. It explains how *E. coli* deals with translation of incomplete mRNAs that might arise by premature transcription termination or cleavage by ribonuclease.

Non-contiguous decoding was known previously in the bypassing of 50 nucleotides in the mRNA for bacteriophage T4 gene 60. Ribosomes translating this mRNA efficiently move from a take-off site to a landing site through a complex set of signals in the mRNA and nascent peptide⁶. In the case of tagging involving 10Sa RNA, the take-off and landing sites are on two different RNA molecules. Because polyuridylic acid works as the take-off RNA (H. Himeno *et al.*), presumably specific signals in the first mRNA are not required, merely a 3' end without a stop codon. Specificity is encoded in the 10Sa RNA. So it seems unlikely that this particular mechanism will be used in a more general sense to allow ribosomes to switch from one mRNA to another. However, the door is open for variations on this theme.

Even without generality, the 10Sa molecule is of considerable interest in evolutionary terms. An evolving mechanism for conversion of an RNA sequence to an amino-acid sequence needs both mRNA and adaptor tRNAs. The 10Sa RNA molecule is provocative in that it embodies both functions. □

John F. Atkins and Raymond F. Gesteland, a Howard Hughes Medical Institute Investigator, are in the Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA.

From mountains to basin

Gregory Houseman

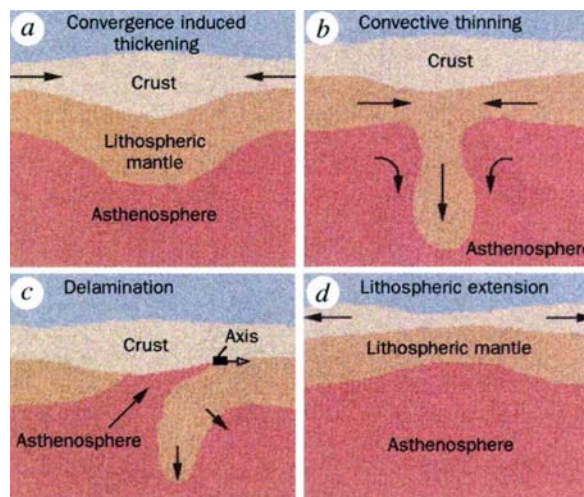
It has been speculated that sometimes sub-continental lithosphere is recycled into the convecting mantle below, but the process has never been clearly seen. On page 785 of this issue¹, Seber and co-workers describe new geophysical evidence for the process of lithospheric delamination occurring beneath the western Mediterranean.

The lithosphere is the outermost part of the Earth, consisting of the crust and a cold, relatively rigid layer of outer mantle. It is usually depicted as a stable layer overlying a vigorously convecting mantle. Although the continental lithosphere is generally exempt from the process of subduction, by which oceanic lithosphere is recycled, its interaction with the mantle is even more interesting because it is harder to detect.

The lower part of the continental lithosphere is colder than the mantle underneath it, and so denser. This configuration creates a potential for gravitational instability that has been envisaged in different ways (parts *b* and *c* in the figure): the unstable layer might peel off, or 'delaminate', and sink into the mantle²; or there may be convective thinning³, in which a previously thickened lithospheric layer is thinned by viscous flow. Convective thinning may occur as a Rayleigh-Taylor instability, by which a layer of dense fluid above less dense fluid overturns⁴. With delamination, as originally defined, the lower layer is peeled away from a migrating 'axis of delamination'. The layer is flexed but not necessarily stretched as it is pulled downwards by gravity. With convective thinning, the locus of downwelling may be fixed and the layer is strongly sheared and stretched as it is pulled down. The term 'delamination' is, however, often loosely used to refer to both processes.

However the instability occurs, lithosphere is replaced by hot, less dense asthenosphere, lifting up the surface and changing the local stress balance. The resulting uplift greatly increases the gravitational potential energy of the region, and so may cause the remaining lithosphere to become extended^{5,6} (part *d* of the figure). The isostatic readjustment that accompanies thinning of the low-density crust then causes surface elevation to fall.

Examples of lithospheric thinning events are perhaps frequent in the geological record⁶. Prominent among recent examples, the Colorado Plateau⁷, Tibet⁵⁻⁷ and the Alboran Sea^{6,8} all show the signs of such events during the Miocene era — between about 7 and 26 million years (Myr) ago. In many cases where a thinning event is inferred, the trigger for instability is a preceding episode of crustal thickening caused by continental collision (part *a* in the figure). Tibet is perhaps the archetype here, because the geological



Stages in the evolution of a thickened region of lithosphere: *a*, thickening produced by horizontal convergence; *b* and *c*, two possible ways of removing sub-crustal lithosphere, by convective thinning or by delamination respectively; *d*, subsequently extended crust above thinned mantle lithosphere.

record clearly shows a relatively sudden switch from north-south convergence to east-west extension around 8 to 10 Myr ago^{3,7}, while India continued to plough steadily northwards at 5 cm yr⁻¹. A lithospheric thinning event is the simplest way to explain the apparent contradiction of extension in the midst of a compressional environment. The Alboran Sea, in the western Mediterranean, is another example of crustal extension occurring in the middle of a continuing continental collision⁸ — in this case between Africa and Spain. The horizontal scale of the process in the Alboran Sea (of the order of 300 km) is much less than in Tibet, but extension of the Alboran Sea crust has been so catastrophic that most of the region now lies below sea level.

The evidence cited above is indirect, in that we see the consequences of the process rather than the process itself. For that reason it is difficult to tell whether lithospheric thinning is more like delamination, where the lower layers of the lithosphere are peeled off coherently², or more like convective thinning, where the

1. Tu, G. F. *et al.* *J. biol. Chem.* **270**, 9322-9326 (1995).
2. Keller, K. C., Waller, P. R. H. & Sauer, R. T. *Science* **271**, 990-993 (1996).
3. Komine, Y. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **91**, 9223-9227 (1994).
4. Ushida, C. *et al.* *Nucleic Acids Res.* **22**, 3392-3396 (1994).
5. Jentsch, S. & Schlenker, S. *Cell* **82**, 881-884 (1995).
6. Weiss, R. B. *et al.* *Cell* **62**, 117-126 (1990).
7. Retallack, D. M. & Friedman, D. I. *Cell* **83**, 227-235 (1995).

layers drip off in a kind of viscous flow embedded in the small-scale convection of the upper mantle^{3,9}. Observations of the overlying crust do not distinguish easily between the two mechanisms because, in either case, cold mantle is replaced by hot mantle and the effect on the crust is similar. But now the increasing capacity of regional seismic networks to locate earthquakes accurately and to measure the properties of the medium traversed by the seismic rays permits us to look more closely at a lithospheric thinning event.

The evidence presented by Seber *et al.*¹ shows extensive seismic activity in a region of the upper mantle beneath the Alboran Sea. Apart from isolated deeper events, most of the activity in this zone appears at depths of between 60 and 150 km. The high seismic activity combined with evidence for increased seismic-wave velocity in the same region shows that this lithospheric layer is relatively cold and brittle. Overlying the high-velocity layer, between depths of about 20 km and 60 km, is a layer of mantle material in which there is no earthquake activity, and which strongly attenuates seismic S waves (transverse waves which only propagate through solids). This layer also has relatively low seismic velocity, so it is interpreted as a zone of hot asthenosphere which has somehow replaced the cold layer of upper mantle.

The authors interpret all this as a snapshot of the delamination process. The high wave-velocity, seismically active layer is the original lithosphere now sinking into the mantle, and the aseismic zone is asthenosphere which has moved upwards to replace the descending lithosphere. One interpretation examined by the authors is of a delamination-style thinning event, in which a northwest-dipping lithospheric slab is sinking, with the axis of delamination migrating in a southeasterly direction (see part *b* in the figure). This interpretation is consistent with reflection-seismic and well data that suggest that subsidence in the Alboran Sea basin propagated in the east/southeast direction¹⁰.

The picture is of course more complex than this simple description permits, as the geological evidence suggests that thinning, extension and thermal metamor-

phism began in the early Miocene before about 22 Myr ago, and extension continued through to the late Miocene⁸ some 8 to 10 Myr ago. The mechanical instability of the thinning event should proceed relatively rapidly once initiated, so that material removed from the lower lithosphere at the time of maximum activity is now perhaps deep in the mantle.

What is now evident of the thermal structure and seismic activity of the upper mantle beneath the Alboran Sea should perhaps be interpreted as a late readjustment to this Miocene event. The presence of high-velocity, seismically active, lithospheric material at depths of 60 to 150 km

beneath the Alboran Sea today does not necessarily fit comfortably with either of the simplified conceptual models. I think this observation is, however, more easily explained in terms of a convective thinning model. The seismicity might be showing the location of a high-strain region in which cold lithosphere now continues to flow down a descending conduit created during the initial convective instability beneath the Alboran Sea. □

Gregory Houseman is in the Departments of Earth Sciences and Mathematics, Monash University, Wellington Road, Clayton, Victoria 3168, Australia.

HUMAN GENETICS

Complex traits on the map

Jurg Ott

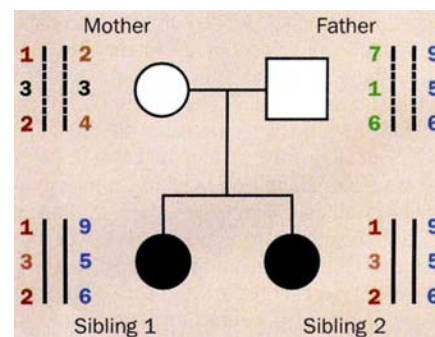
NOWADAYS it is fairly straightforward to localize genes responsible for diseases that have a simple genetic basis — that is, diseases following a clear mendelian mode of inheritance, such as cystic fibrosis or Huntington's disease. All one needs to do is carry out a 'genome screen' by typing individuals in suitable families for a good number (150 or more) of genetic markers spread over the human genome (the chromosomes). But it is much more difficult to do this for so-called complex traits, schizophrenia for example, that are characterized by non-mendelian inheritance and other disturbing features such as equivocal disease classification.

Hugot and colleagues¹ (page 821 of this issue) have successfully tackled such a problem, and their approach can be taken as a model case. Using state-of-the-art genetic analysis techniques, they have mapped one of the (presumably several) genes responsible for Crohn's disease (CD) to an area on chromosome 16 near its centromere. Crohn's disease results in chronic inflammation of the bowel, and it has long been thought that there are genetic components that predispose sufferers to this and similar conditions; as compared with the general population, first-degree relatives of patients with CD are ten times more likely to contract the disease².

The initial steps for localizing a disease gene are more or less the same for simple and complex traits, except that the latter require much larger numbers of families to be studied. With an initial set of 25 families, Hugot and co-workers genotyped family members for 270 highly polymorphic markers spread over the genome. For each pair of loci (hypothesized disease locus and a given marker locus), linkage analysis^{3,4} showed whether alleles at the two loci had a tendency to be co-

inherited when passed from parents to offspring, which would indicate close proximity between the two loci.

The statistical 'currency' of evidence for linkage between a disease locus and marker is the lod score. This is the logarithm of the ratio of two pedigree data likelihoods — the one obtained assuming linkage, divided by the one obtained under no linkage (a lod score of 3 or higher is considered strong evidence for linkage). The maximum lod score obtained



Discriminating between alleles in a homozygous genotype (3/3, mother) through analysis of closely linked markers. It is virtually certain that the '3' on the left was passed to each child as each received the '1' and '2' alleles on one chromatid.

for CD occurred with marker D16S409 and was equal to 2.04. Previously, using computer simulation, Hugot and co-workers had carried out power calculations, which predicted with a power of over 97 per cent that a lod score of 3 or higher should be obtained with the given family data; the simulated expected (average) lod score is not given in the paper but must have been much higher than 3. The large discrepancy between the expected lod score and that actually obtained is a good indication that the mode of inheritance of CD is more complex than simple

- Seber, D., Barazangi, M., Ibenbrahim, A. & Demnati, A. *Nature* **379**, 785–790 (1996).
- Bird, P. *J. geophys. Res.* **84**, 7561–7571 (1979).
- Houseman, G. A., McKenzie, D. P. & Molnar, P. *J. geophys. Res.* **86**, 6115–6132 (1981).
- Whitehead, J. A. Jr & Luther, D. S. *J. geophys. Res.* **80**, 705–717 (1975).
- England, P. C. & Houseman, G. A. *Phil. Trans. R. Soc. A* **326**, 301–320 (1988).
- Platt, J. P. & England, P. C. *Am. J. Sci.* **294**, 307–336 (1994).
- Molnar, P., England, P. C. & Martinod, J. *Rev. Geophys.* **31**, 357–396 (1993).
- Platt, J. P. & Vissers, R. L. M. *Geology* **17**, 540–543 (1989).
- Lenardic, A. & Kaula, W. M. *J. geophys. Res.* **100**, 15193–15203 (1995).
- Docherty, C. & Banda, E. *Tectonics* **14**, 804–818 (1995).

recessive, as assumed in the analyses.

For two-point linkage analysis between disease locus and a given marker, two approaches can be taken — parametric, in the form of the standard lod score analysis, or non-parametric, for instance the affected sibpair (ASP) analysis. In the parametric approach, the recombination fraction (equivalent to genetic distance when small) between disease and marker loci is estimated on the basis of the family data and the mode of inheritance (dominant or recessive, with given penetrances) assumed for the trait. If that is incorrect, the resulting loss of information is not necessarily large. The main effect of a misspecification of mode of inheritance is an overestimation of the recombination fraction, which is then viewed as a nuisance parameter rather than a good indicator of disease locus position.

In ASP analyses, attention is focused on pairs of affected siblings and all linkage information is gained from the inheritance of marker alleles by the two siblings. So no assumptions are made as to the mode of trait inheritance. Specifically, one determines the number of alleles (0, 1 or 2) inherited by sibling 2 that are copies of the same parental alleles as those inherited by sib 1 (the number of alleles shared that are identical by descent, i.b.d.). For example, at the top locus in the figure, sib 2 shares two alleles (1 and 9) i.b.d. with sib 1. With random assortment, the proportion of 0, 1 or 2 alleles shared i.b.d. are expected to be 1/4, 1/2 and 1/4, with a shift to higher numbers when linkage is present. Although the two approaches, parametric and non-parametric, are often considered to be fundamentally different, they have been shown under certain conditions to be completely equivalent^{5,6}.

In multipoint analyses, particularly for complex traits, parametric and non-parametric linkage analyses exhibit very different properties, which is in sharp contrast to the situation for two-point analysis. Because of upward biases in estimates of recombination fraction expected for complex-trait loci, a parametric lod score analysis involving multiple markers over a chromosome tends to estimate a disease locus away from each marker — that is, off the chromosome, even though it may be located well inside the marker map. No such detrimental effects occur with multipoint ASP analysis⁷ (implemented in the MAPMAKER/SIBS computer program), which for the first time has been applied to a genomic screen for complex-trait loci by Hugot and co-workers.

The key to the multipoint ASP approach is that multilocus analysis refers to marker loci only and does not involve the disease locus (except through the fact that the analysis is carried out on affected individuals). Essentially, at a given point in the genome, information from all markers on a chromosome is gathered to esti-

mate i.b.d. sharing in ASPs at that point. For example, at the middle locus in the figure, the two maternal alleles (3/3) are indistinguishable. As recombinations among closely linked markers are rare, it appears likely (perhaps supported by additional offspring not shown in the figure) that alleles at the three marker loci are located on individual chromatids, as shown. Thus, both offspring received the sequence of alleles (the haplotype) 1-3-2 from the mother. When these markers are tightly linked, information from neighbouring loci allows discrimination between the two '3' alleles in the mother: it is virtually certain that the '3' allele on the left was passed to each of the two offspring. In this manner, multimarker analysis has made the middle marker essentially 100 per cent informative in this mating.

Multipoint ASP analysis⁷ is ideal for localizing complex trait loci. Even though the number of families available to Hugot and co-workers was relatively small (78 in the final analysis), this new approach allowed them to localize a gene for CD

with greater confidence than has been possible with conventional methods. Their approach lends itself to finding other genes for CD and those underlying complex traits like the psychiatric diseases. Candidate genes in this area can now be scrutinized for involvement in the pathogenesis of CD. Also, geneticists may apply an array of techniques to find stretches of DNA or single nucleotides that are altered in patients compared with unaffected control individuals. □

Jurg Ott is in the Department of Psychiatry, Columbia University, 722 West 168 Street, New York, New York 10032, USA.

1. Hugot, J.-P. *et al.* *Nature* **379**, 821–823 (1996).
2. Orholm, M. *et al.* *New Engl. J. Med.* **324**, 84–88 (1991).
3. Ott, J. *Analysis of Human Genetic Linkage* (Johns Hopkins Univ. Press, Baltimore, 1991).
4. Terwilliger, J. D. & Ott, J. *Handbook of Human Genetic Linkage* (Johns Hopkins Univ. Press, Baltimore, 1994).
5. Hyer, R. N. *et al.* *Am. J. hum. Genet.* **48**, 243–257 (1991).
6. Knapp, M., Seuchter, S. A. & Baur, M. P. *Hum. Hered.* **44**, 44–51 (1994).
7. Kruglyak, L. & Lander, E. S. *Am. J. hum. Genet.* **57**, 439–454 (1995).

CONDENSED MATTER

On the points of melting

David G. Grier

ALTHOUGH melting and freezing are two of the most common structural phase transitions in everyday experience, surprisingly little is known about their microscopic mechanisms. Part of the difficulty in crafting a complete understanding of such fundamental transformations is that experiments on conventional materials, such as copper or water, can tell us very little about what happens on the atomic scale as a crystal transforms itself into a fluid. Not only are atoms small and rapidly moving, but comparatively few of them take part in the process at any one time. Even computer simulations are hard-pressed to accommodate the enormous sample sizes and the range of timescales involved in converting a chunk of ice into a puddle of water. The plasma crystal described by Thomas and Morfill on page 806 of this issue¹ is a new model system in which some of the mysteries of structural phase transitions can be played out on a stage that scientists can watch with ordinary video cameras.

The plasma crystal works on the straightforward principle that like-charged particles repel each other. The particles in Thomas and Morfill's experiments are polymer spheres about 7 µm across (roughly a tenth of the diameter of a human hair) which acquire their charges through immersion in an ionized gas known as a plasma. Because their container prevents them from moving apart

indefinitely, the spheres adopt a configuration which minimizes their total energy, given their temperature and density. If the charge-mediated interaction between the spheres is great enough to overcome the buffeting of the surrounding gas, then they form into regularly spaced arrays like the orderly ranks of atoms in crystals. If the thermal energy of the gas wins, the plasma crystal melts into a dynamic and disordered state reminiscent of a fluid.

Investigations of other model systems have been made in recent decades². Microscopic spheres colloiddally suspended in a fluid solvent also undergo a variety of phase transitions depending on their density, chemical environment and temperature. Even so-called hard spheres, which interact only on contact, undergo a freezing transition to a crystalline state. Colloidal spheres with charged groups chemically grafted to their surfaces can form crystals of different symmetries as well as fluids and glasses. Mixing spheres of two or more different sizes leads to even greater complexity which has yet to be explored fully.

Thermal equilibrium in a conventional material's ensemble of atoms is maintained by interatomic collisions. These are energy- and momentum-conserving processes mediated by interactions with strong quantum mechanical contributions. In a colloidal suspension, the temperature is set by the surrounding fluid, and viscous

drag means that energy and momentum need not be conserved in collisions, so there are major differences in the microscopic dynamics underlying colloidal and conventional phase transitions.

The plasma in a plasma crystal is maintained by a radio-frequency discharge which deposits energy into and partially ionizes a low pressure gas. The radio-frequency source also exerts forces on the charged microspheres, and the resulting excess energy is transferred through collisions to the background of neutral gas molecules at roughly room temperature. Lowering the pressure of the buffer gas lowers the rate at which energy is removed from the ensemble of spheres and thus allows its equilibrium temperature to rise. As in colloidal suspensions, thermal equilibrium is maintained by collisions between the spheres. Unlike fluid suspensions, however, collisions in a plasma crystal should be very nearly elastic, because the low-pressure gas provides comparatively little viscous damping. By adjusting the pressure of buffer gas, Thomas and Morfill can effectively control the temperature of the ensemble of spheres and watch the plasma crystal melt at 'atomic' resolution.

Some aspects of the melting process they see are familiar to everyone; others have been discussed in the technical literature; still others appear strange and novel. The plasma crystal consists of a stack of a dozen or so horizontal layers with charged spheres arranged in lattices that have six-fold symmetry in each layer. The particles vibrate about their lattice positions much as thermal phonons set atoms vibrating in conventional materials. The crystal is subject to the same types of topological lattice defect that crop up in conventional and colloidal crystals.

As Thomas and Morfill raise the effective temperature, the plasma crystal melts. Near the melting transition, residual crystallites coexist with the resulting fluid much like ice cubes floating in water. The fluid at this stage has unusual properties. Rather than engaging in random brownian motion like atoms in conventional fluids at equilibrium, the 'plasma fluid' has strongly correlated flows within the plane and little exchange of particles between planes. Layering in dense fluids is well known in colloidal suspensions^{3,4} and has been observed using the surface force apparatus in conventional fluids⁵. If the strongly correlated planar motion is the plasma crystal equivalent of conventional convection, how is it driven? If not, then what is it? Answering these questions will require a deeper understanding of the local energetics of plasma crystals than is currently available.

When they raise the temperature further, the "flow and floe" morphology gives way to a uniformly disordered fluid which retains a degree of layering and has a sur-

prising amount of orientational order in the plane, the particles arranging themselves into patterns with local six-fold rotational symmetry. This strange state of matter has been likened to the orientationally ordered 'hexatic' fluid, although the hexatic phase is predicted to occur only in two dimensions, not three⁶. Although the oriented fluid state might be peculiar to the plasma crystal, it could be a sign of a new intermediate stage of melting in a wide range of materials.

Finally, at the lowest gas pressures and highest temperatures, the oriented, layered fluid crosses over to a more familiar isotropic homogeneous gas-like phase.

Plasma crystals also display intriguing non-thermal properties. Rather than jostling randomly, the spheres trace out elliptical trajectories under some conditions. Presumably a response to the radio-frequency driving field, these motions may provide a means to measure the interactions among the spheres as well as between the spheres and the field.

Studying this interplay may lead to an understanding of the startling non-uniformities in density observed in the fluid

under some conditions. Are the observed fluctuations examples of equilibrium phase separation or do they represent a crossover to non-equilibrium behaviour?

Answers to such questions will deepen our understanding of this exciting new state of matter. They will also enable researchers to apply results from plasma crystal experiments to resolving fundamental problems in condensed matter physics and revealing the mechanisms and pathways of structural phase transitions. □

David G. Grier is in the Department of Physics and at the James Franck Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA.

1. Thomas, H. M. & Morfill, G. E. *Nature* **379**, 806–809 (1996).
2. Murray, C. A. in *Bond-Orientational Order in Condensed Matter Systems* (ed. Strandburg, K. A.) 137–215 (Springer, New York, 1992).
3. Murray, C. A., Sprenger, W. O. & Wenk, R. A. *J. Phys.: Condens. Matter* **2**, SA385–SA388 (1990).
4. Grier, D. G. & Murray, C. A. *J. chem. Phys.* **100**, 9088–9095 (1994).
5. Israelachvili, J. N. & McGuigan, P. M. *Science* **241**, 795–800 (1988).
6. Nelson, D. R. in *Phase Transitions and Critical Phenomena* Vol. 7 (eds Domb, C. & Leibowitz, J. L.) 1–99 (Academic, London, 1983).

ASTRONOMY

A new rapid X-ray repeater

Christopher Thompson

A TYPICAL cosmic X-ray burst releases more radiant energy in ten seconds than the Sun does in a day. First detected in 1975 (refs 1 and 2), most of these bursts have been convincingly explained as thermonuclear flashes on the surfaces of neutron stars³, but a few sources have held on to their secrets more stubbornly. These include the Rapid Burster, discovered⁴ early in 1976, which releases staccato flashes of X-rays as often as once every 7 seconds, and the even stranger triplet of soft gamma repeaters⁵, whose much rarer bursts outshine a typical X-ray burst by a factor of ten thousand. Now, after a long wait, the Compton Gamma Ray Observatory has discovered a second rapidly repeating X-ray burster, since dubbed GROJ1744–28. The discovery is reported by Kouveliotou and colleagues on page 799 of this issue⁶.

This object sits in our Galaxy along with almost all the repeating X-ray sources. (One soft gamma repeater appears to be in a satellite galaxy, the Large Magellanic Cloud.) By contrast, classical, non-repeating gamma-ray bursts are probably much more violent events at cosmological distances⁷. The new source, which lies in the direction of the Galactic Centre, was first seen on 2 December. It emits bursts lasting between 6 and 100 seconds and arriving at intervals that have increased from about 200 seconds to about an hour.

To appreciate the potential significance of GROJ1744–28, let us review some of the mysteries of the Rapid Burster. Along with some thermonuclear flashes, it emits a large number of bursts that involve a distinctly new mechanism, and are called type-II bursts. They release more energy than can be accounted for by nuclear fusion of the fuel accumulated between bursts (at most 8.4 MeV per nucleon, burning hydrogen to iron), but just the amount expected from infall onto the surface of a neutron star (200 MeV per nucleon). Kouveliotou *et al.* deduce that the bursts of GROJ1744–28 cannot be powered by nuclear burning from an observational upper limit on the X-ray flux in between bursts, which limits the rate of fuel intake. The rapid and almost regular repetitions then imply that GROJ1744–28 is a neutron star undergoing spasmodic accretion.

How type-II X-ray bursts are triggered remains something of a mystery. The magnetic field of the neutron star is probably involved, acting as a gate that holds back some material from the accretion flow until a critical mass is reached. This gate appears to operate with less efficiency and regularity in GROJ1744–28 than it does in the Rapid Burster.

Given that GROJ1744–28 is probably a neutron star, how does the strength of its

magnetic field compare with other burst sources? The X-ray photons it emits are higher in energy than those emitted by the Rapid Burster, which leads Kouveliotou *et al.* to suggest that GROJ1744–28 has a stronger magnetic field. Support for this inference comes from the detection of coherently pulsed, persistent X-ray emission^{8–10}. The 0.467-second pulse period is almost certainly the rotation period of the star.

Conventional wisdom holds that neutron stars — including the Rapid Burster — that emit type-I (thermonuclear) flashes are weakly magnetized, with surface fields of the order of 10^8 to 10^9 gauss; whereas pulsating X-ray sources have stronger magnetic fields of the order of 10^{11} to 10^{13} gauss which channel the accreted material to polar hotspots, thereby raising the temperature above the threshold for stable nuclear burning³. The absence of coherent X-ray pulsations in type-I burst sources is ascribed to the weakness of the magnetic field. This reasoning has always been problematic in the case of the Rapid Burster, for the magnetic field of that source is apparently strong enough to trigger type-II X-ray bursts as well. Some additional ingredient may be required to suppress pulsations, such as alignment of the magnetic axis with the angular momentum of the accreting matter.

GROJ1744–28 may already be giving us some clues about that most basic of neutron-star puzzles, namely how their magnetic fields decay. The tendency of weakly magnetized radio pulsars to spin rapidly and to have binary stellar companions has led to a growing suspicion, not as yet backed by any compelling physical theory, that deposition of enough matter on the surface of a neutron star can bury its magnetic field. If the Rapid Burster is indeed weakly magnetized, then it is surprising that its most energetic bursts should occur at similar intervals (a few hundred seconds to one hour) and with similar energies to the bursts from GROJ1744–28.

But there is a loophole. The idea of field burial allows for the existence of neutron stars which have experienced only partial burial: that is, stars that are weakly magnetized over most of their surface and can emit thermonuclear flashes, but still retain a sizeable fraction of the original external dipole flux. The behaviour of type-II bursts suggests the existence of two distinct accretion channels¹¹ which may reflect the distribution of magnetic flux over the stellar surface.

How do the soft gamma repeaters (SGRs) fit in with these two rapid X-ray repeaters? SGRs only emit a faint glow of X-rays in between their brief, blinding outbursts, during which their X-ray output increases by a factor of a million or more. By contrast, the Rapid Burster only

brightens by an order of magnitude when it bursts. This indicates that SGR bursts are probably not powered by accretion instabilities. Thermonuclear power has great difficulty accommodating the extraordinary burst emitted by SGR0526–66 on 5 March 1979, which released a million times the energy of a typical type-I X-ray burst. One possibility that has received considerable recent attention¹² is that the SGR sources are a rare subspecies of neutron stars with magnetic fields stronger than 10^{14} gauss, and that bursts are triggered when these very strong fields fracture their crusts.

This does not mean that there are no threads connecting the SGR sources with the rapid X-ray repeaters. Curiously, the spectra of SGR bursts and the Rapid Burster's type-II bursts show a number of similarities. Both are consistent, to a first approximation, with black-body emission from an area comparable to the surface of a neutron star. (SGR bursts are much more luminous and correspondingly much hotter and spectrally harder.) Their radiating surfaces appear to contract at approximately constant temperature, without the marked spectral softening (and occasional photospheric expansion) characteristic of type-I X-ray bursts. Together these observations suggest that the emitting plasma is optically thick and confined by the magnetic field of the neutron star. One explanation for this similarity of spectral behaviour in such otherwise dissimilar sources is that the free energy is absorbed by the magnetic field and converted to a trapped photon–electron–positron plasma¹².

It is not yet clear whether the bursts of GROJ1744–28 share this radiative mechanism. They are harder than the type-II bursts emitted by the Rapid Burster and much less luminous than SGR bursts. It may be that X-rays are emitted from a small patch of the star's surface. In any case, the message flashed to us from the universe of X-ray astronomy is a subtle one, of both unity and diversity. □

Christopher Thompson is in the Department of Physics and Astronomy, University of North Carolina, Chapel Hill, North Carolina 27599–3255, USA.

- Grindlay, J. E. *et al. Astrophys. J.* **205**, L127–L130 (1976).
- Bellian, R. D., Conner, J. P. & Evans, W. D. *Astrophys. J.* **206**, L135–L138 (1976).
- Lewin, W. H. G., Van Paradijs, J. & Taam, R. E. *Space Sci. Rev.* **62**, 223–389 (1993).
- Lewin, W. H. G. *et al. Astrophys. J.* **207**, L95–L99 (1976).
- Norris, J. P. *et al. Astrophys. J.* **366**, 240–252 (1991).
- Kouveliotou, C. *et al. Nature* **379**, 799–801 (1996).
- Kouveliotou, C., Briggs, M. S. & Fishman, G. J. (eds) *Proc. 3rd Huntsville Gamma Ray Burst Workshop* (Am. Inst. Phys., in the press).
- Finger, M. H. *et al. IAU Circ. No.* 6285 (1996).
- Kouveliotou, C. *et al. IAU Circ. No.* 6286 (1996).
- Swank, J. *et al. IAU Circ. No.* 6291 (1996).
- Dotani, T. *et al. Astrophys. J.* **350**, 395–407 (1990).
- Thompson, C. & Duncan, R. C. *Mon. Not. R. astr. Soc.* **275**, 255–300 (1995).

Per aqua ad astra

Guns of various kinds have often been proposed as space launchers, but have never gained wide acceptance. Daedalus has a new approach to the topic.

The ocean, he points out, exerts an atmosphere of pressure for every 10 m of depth, and in some places is over 10 km deep. Imagine, he says, a tube extending right to the bottom, with a free-rising piston at its base. This is a gun with an initial barrel pressure of 1,000 atmospheres and a length of 10 km. A 1-m diameter tube could release enough energy to accelerate a 6 tonne payload to the escape velocity of 11 km s^{-1} .

Sadly, it would never do it. Tons of water would be accelerated too, wasting most of the energy; air could not be expelled fast enough from the long barrel, and no payload could rise through the sea-level atmosphere at 11 km s^{-1} . Daedalus sees the idea not as a replacement for a complete space rocket, but merely for its big first stage. This accounts for about 80% of the weight and cost of a space launcher, but gives it only about 20% of its final velocity.

The oceanic catapult will consist of many short segments of steel pipe held together by O-ring seals, each with a small solid-charge actuator to separate the joint on command. As the piston rises up the long tube, each joint is separated the instant it has passed. The segment falls away, and the water rushes sideways into the upwardly disintegrating barrel. None needs to be accelerated upwards.

A cushioning slug of compressed gas above the entering water will transmit its pressure to the piston ahead of it, on which the vehicle rests. The barrel above it contains no air to be pushed out of the way. It is full of vacuum — one extra atmosphere of compressive stress hardly matters in such a design. It is sealed at the top by a lid which is blown open just before the vehicle emerges at a few times the speed of sound: just about the velocity which a conventional first stage would have given it. The second and subsequent rocket stages will then take over to complete the launch. A catapult 1 m across could launch a rocket of several hundred tonnes.

Much Shuttle hardware is already designed to be fished out of the ocean and used again, and deep-sea oil drillers are already expert at assembling pipes under water. So the oceanic launcher might not need much development. Once built, it could be used repeatedly. Its barrel segments would be retained on cables, to be hauled together and pumped out ready for the next launch. Engineers would applaud its simplicity, scientists its cheapness, and environmentalists its lack of pollution.

David Jones

Pulling the trigger on psoriasis

SIR — Psoriasis is a common inflammatory skin disease, possibly induced by a T-cell-mediated autoimmune reaction triggered by bacterial superantigens¹. Some transgenic animal models² and mutations in mice³ mimic certain aspects of psoriasis, but they do not allow study of the role of components of the immune system or exogenous stimuli, namely T cells and superantigens, in the onset of psoriatic lesions. Thus, we searched for a model in which a psoriasiform inflammation could be triggered.

Eighteen full-thickness skin grafts were transplanted from clinically uninvolved areas of three psoriatic patients onto mice with severe combined immunodeficiency (SCID), lacking B and T cells as described previously⁴. Technical manipulations (transplantation, injection) did not alter the state of the grafts as seen in controls treated with repetitive intradermal saline injections. Grafts injected with 3 µg of the bacterial superantigen exfoliative toxin showed some of the hallmarks of psoriasis (see table); these included profound epidermal thickening (akanthosis) from

hyperproliferating basal keratinocytes, documented by expression of the mitosis-associated Ki-67 antigen, along with increased indentation of epidermis and dermis (papillomatosis) and focal neo-expression of the adhesion molecule ICAM-1 on basal and suprabasal keratinocytes overlying the papillary dermis. This pattern is thought to be characteristic of psoriasis⁵.

When the patients' superantigen-stimulated peripheral blood mononuclear cells were simultaneously injected intraperitoneally, an epidermotropic T-cell infiltrate, positive for the cutaneous lymphocyte-associated antigen, was observed (see table). Marked T-cell epidermotropism was an exclusive feature of grafts from psoriatic donors simultaneously treated with exfoliative toxin and peripheral blood mononuclear cells. None of these phenomena was noted in 18 transplants from 3 non-psoriatic controls.

Our observations document the potential of superantigens to trigger psoriasis in a conditioned environment. This experimental approach should greatly help studies on the pathogenesis of psoriasis. In contrast to

other models described so far, it allows the specific induction of psoriasiform inflammation. Moreover, this model system provides an excellent means of studying the modes of superantigen-induced (auto-) immune responses, as well as the phenomenon of lymphocyte homing in general, and epidermotropism in particular.

W.-H. Boehncke

D. Dressel

Department of Dermatology,
University of Ulm, 89081 Ulm, Germany

T. M. Zollner

R. Kaufmann

Department of Dermatology,
University of Frankfurt,
60590 Frankfurt, Germany

1. Valdimarsson, H., Baker, B. S., Jonsdottir, I., Powles, A. & Fry, L. *Immun. Today* **16**, 145–149 (1995).
2. Wilson, J. B., Weinberg, W., Johnson, R., Yuspa, S. & Levine, A. J. *Cell* **61**, 1315–1327 (1990).
3. Sundberg, J. P., Dunstan, R. W., Roop, D. R. & Beamer, W. G. *J. invest. Derm.* **102**, 781–788 (1994).
4. Boehncke, W.-H., Sterry, W., Hainzl, A., Scheffold, W. & Kaufmann, R. *Archs derm. Res.* **286**, 325–330 (1994).
5. Nickoloff, B. J., Griffith, C. E. & Barker, J. N. W. N. *J. invest. Derm.* **94**, 151S–157S (1990).

RIP for viruses

SIR — Many plants contain ribosome-inactivating proteins (RIPs) with *N*-glycosidase activity, which depurinate large ribosomal RNA from sensitive ribosomes, thus arresting protein synthesis¹. All RIPs so far tested inhibit the replication of plant and animal viruses². The exact role of these proteins in plants is unclear, but it has been proposed that they exert a defensive, antiviral function³. We report here that viral infection induces the expression of two single-chain RIPs (beetin 27 and beetin 29) in sugar beet (*Beta vulgaris*). A similar expression was obtained by treating beet leaves with hydrogen peroxide or salicylic acid, two mediators of plant acquired resistance⁴.

Two proteins, beetin 27 (relative molecular mass 27,000) and beetin 29 (*M*_r 29,000), were isolated from leaves of sugar beet infected with beet mild yellowing virus, by a procedure used to purify single-chain RIPs⁵. Beetins displayed several properties of RIPs: beetin 27 (the more abundant) inhibited cell-free protein synthesis and promoted the release of the

MORPHOLOGICAL COMPARISON OF GRAFTS FROM NORMAL VOLUNTEERS AND UNINVOLVED SKIN FROM PSORIATIC PATIENTS

Source*	Normal	Normal	Normal	Psoriatic	Psoriatic	Psoriatic
Treatment	PBS	ET	ET/PBMCs	PBS	ET	ET/PBMCs
Epidermal thickness†	111 (90/123)	140 (120/168)	143 (113/166)	149 (143/165)	306 (300/417)	327 (262/450)
Papillomatosis index‡	1.2 (1.2/1.3)	1.3 (1.2/1.3)	1.2 (1.2/1.3)	1.3 (1.2/1.3)	2.2 (1.9/2.4)	2.1 (2.0/2.3)
Ki-67§	Negative	Single cells	Single cells	Single cells	Positive	Positive
ICAM-1	Negative	Negative	Negative	Negative	Focal	Focal
T cells	Few	Few	Few	Few	Few	Epidermotropic¶

Transplantations were done as described previously⁴. Six full-thickness skin grafts were prepared from the inner aspect of the forearm of each of three healthy volunteers and from clinically uninvolved skin at the same location from three psoriatic patients. Two grafts of each individual received a similar treatment: intradermal injection with PBS; with ET diluted in PBS; or with ET diluted in PBS and simultaneous intraperitoneal injection of the donor's PBMCs. Intradermal injections consisted of 100 µl PBS with or without 3 µg ET (Toxin Technology, Florida). Intraperitoneal injections contained 2 × 10⁶ of the donor's PBMCs in 100 µl PBS. PBMCs were stimulated in supplemented RPMI 1640 medium containing 100 ng ml⁻¹ ET 48 h before injection at a density of 1 × 10⁶ cells per ml. Injections were given at days 28, 31, 34 and 37 after transplantation; mice were killed at day 40. Grafts were handled for immunoperoxidase and immunofluorescence staining as reported previously⁴. ET, exfoliative toxin; PBMCs, peripheral blood mononuclear cells; PBS, phosphate-buffered saline.

*Each cohort consists of six grafts from three different individuals. Experiments were done one individual at a time, and so this table summarizes the data from six independent experiments.

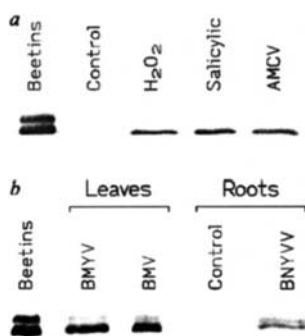
†Median (minimum/maximum), indicated in µm.

‡Length ratio of dermo-epidermal border to skin surface; medians (minimum/maximum) are indicated.

§Staining pattern of basal keratinocytes.

||Staining was observed on keratinocytes overlying the papillary dermis only.

¶Most intra-epidermal CD3⁺ cells were also positive for the cutaneous lymphocyte-associated antigen.



Western blot immunoblot analysis of the induction of beetins in leaves of sugar beet plants by mediators of systemic acquired resistance and viral infection. Purified beetins prepared as described elsewhere⁵, or protein extracts from either hydrogen peroxide- or salicylic acid-treated plants, or virus-infected plants, were separated by SDS-PAGE and electroblotted onto Immobilon-P membranes. Beetins were detected as described elsewhere⁹. *a*, Protein extracts from laboratory-grown plants; *b*, protein extracts from field-grown plants. Control, uninfected plants; AMCV, artichoke mottle crinkled virus; BMYV, beet mild yellowing virus; BMV, beet mosaic virus; BNYVV, beet necrotic yellow vein virus.

rRNA fragment typical of the single depurination operated by RIPs¹. The amino-terminal amino-acid sequence of both proteins had 45% amino-acid identity with RIPs possessing anti-HIV-1 activity¹ (data not shown). Beetins may also promote direct depurination of genomic tobacco mosaic virus RNA (data not shown), but the correct physiological relevance of such hydrolytic reactions is unknown. Western blot analysis with rabbit anti-beetin antibodies revealed that beetins were present in leaves of plants infected in the laboratory with artichoke mottle crinkled virus, leaves of field-grown

beet infected with beet mosaic virus, and the roots of plants field-infected with beet necrotic yellow vein virus, whereas they were not detectable in extracts from leaves and roots of healthy plants (see figure).

Using the same technique, both beetins were detected in extracts from healthy, laboratory-grown plants treated topically with diluted solutions of salicylic acid or H_2O_2 (see figure), whose concentration increases transiently in plants after viral infections and which are considered to be mediators of the virally induced acquired resistance in host plants. Laboratory-infected plants had no severe disease.

Viral proliferation can be arrested by: a sort of 'suicide action', through the inactivation of homologous ribosomes with consequent cell death⁶; through a direct effect on viral RNA⁷; by an effect on virally derived DNA, because at least some RIPs depurinate DNA⁷; or by a combination of these and other unknown mechanisms. Our results show that viral infection and mediators of acquired resistance induce the expression of RIPs in plant tissues. This supports the proposal that these proteins have an antiviral role, giving an evolutionary advantage to the plant, a notion supported by the resistance to viral infections of transfectants plants expressing *Phytolacca americana* RIP⁸.

**Tomás Gírbés, Carlos de Torre
Rosario Iglesias, José Miguel Ferreras**

Departamento de Bioquímica
y Biología Molecular,
Facultad de Ciencias,
Universidad de Valladolid,
47005 Valladolid, Spain
Enrique Méndez
Servicio de Endocrinología,
Centro Ramón y Cajal,
Carretera de Colmenar Viejo km 9,
28034 Madrid, Spain

Dinosaur nesting or preying?

SIR — We congratulate Norell *et al.*¹ on their discovery of a nesting dinosaur. We would like to make a suggestion to support the view that the animal was indeed incubating a clutch of eggs. There is a reasonable chance that the dinosaur was female. In most extant birds, females incubate the eggs, in many species both males and females incubate, and in only a few species do males exclusively incubate².

Female birds produce medullary bone as a calcium phosphate reservoir before egg laying that is laid down rapidly, and

likewise resorbed rapidly. It has a very characteristic fabric, and can easily be identified in sectioned bone³. It would be most interesting to learn if a similar tissue type is present in the long bones of the nesting *Oviraptor*. The presence of such a tissue type would be convincing evidence that the animal was female and that it was indeed in its oestrous cycle, and would also confirm its relationship to birds.

**David M. Martill
Michael J. Barker
Chris G. Dacke**

Palaeobiology Research Group,
Department of Geology, and School of
Pharmacy and Biomedical Science,
University of Portsmouth
Portsmouth PO1 3QL, UK

- Norell, M. A. *et al.* *Nature* **378**, 774–776 (1995).
- Bentley, P. J. *Comparative Vertebrate Endocrinology* (Cambridge Univ. Press, New York, 1982).
- Dacke, C. G. *et al.* *J. exp. Biol.* **184**, 63–88 (1993).

The load of the Taung child

SIR — Following our recent suggestion that a large bird of prey collected the Taung child and associated fauna¹, Hedenström² suggested that the load-lifting capacity of a crowned eagle, one of several candidate species of African raptor that may have acted as the accumulator, would have been inadequate for lifting the estimated mass of the Taung infant. He therefore suggested that the Taung child must have been dismembered before transport to the nest.

In support of his idea, Hedenström cited a probable weight for the Taung child as 10–12 kg, whereas we suggested that this was its maximum weight. The actual body mass of the Taung child may have been less than 10 kg. We used the mass range cited by Hedenström to demonstrate that the maximum estimated body mass of the Taung child was within the collecting capabilities of several large African raptors. Furthermore, it is common for large raptors to eviscerate large prey before lifting, and mammals may lose as much as 30% of their body weight after being degutted³. Thus, after being eviscerated, even a 12-kg Taung child would be brought near Hedenström's maximum short-range carrying mass (6.1 kg) of an average-sized crowned eagle, and most probably within the short-range carrying capabilities of a large member of this species.

Further, we did not suggest that the Taung child lived in a savannah habitat, but rather that the immediate habitat of Taung during the period of deposition in question (around 2.6 million years ago) was tropical and wooded. We therefore cannot be certain whether the Taung child was carried during an eagle's "short anaerobic sprint" from the surrounding area, or whether it was carried over a "considerable distance".

Both objective and anecdotal observations on the short- and long-range prey-carrying capacities of extant African raptors, including the crowned eagle, clearly demonstrate greater mass-lifting capabilities than the 6.1 and 1.7 kg for short- and long-distance carrying capacities calculated by Hedenström. For example, the black eagle, *Aquila verreauxi* (an eagle on average approximately the same mass, 4.3 kg, as the crowned eagle), has been shown to carry

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.

- Berger, L. R. & Clarke, R. J. *J. hum. Evol.* **29**, 275–299 (1995).
- Hedenström, A. *Nature* **378**, 670 (1995).
- Gargett, V. *The Black Eagle* (Acorn, Johannesburg, 1990).
- Visser, J. *Afr. wild Life* **17**, 191–194 (1963).
- Steyn, P. *Birds of Prey of Southern Africa* (David Phillip, Cape Town, 1982).
- Rowe, E. G. *Ibis* **89**, 387–410 (1947).
- Gargett, V. *Ostrich* **53**, 147–150 (1982).
- Tarboton, W. R. & Allan, D. G. *The Status and Conservation of Birds of Prey in the Transvaal* (Transvaal Mus. Monogr., Pretoria, 1984).
- Jubb, R. A. *Honey Guide* **114/115**, 60 (1983).

weights of killed prey of up to 4.5 kg (ref. 3). It is not known how close to the nest these kills were made, as none was made within sight of the observer. But many studies of large African raptors include subjective assessments of the weights of animals collected and carried by birds over apparently long distances as being greater than the 1.7 kg suggested by Hedenström^{1,3,4-8}. There are also accounts of eagles lifting animals presumably greater than 6 kg over short distances, and remains of animals found below nests suggest that this is not an uncommon occurrence^{1,3,5,8,9}.

Finally, as we have previously noted¹, we cannot discount the possibility that an extinct bird of prey, larger in body size than either the extant crowned or black eagle, was responsible for the fossil accumulations at Taung.

Lee R. Berger

Ronald J. Clarke

*Palaeo-Anthropology Research Unit,
Department of Anatomical Sciences,
Medical School,
University of the Witwatersrand,
7 York Rd, Parktown,
Johannesburg 2193, South Africa*

evolution of dioecy in this plant family, by providing an additional advantage to individuals that lose their female function if they concomitantly acquire the ability to shed flowers early and hence reduce their risk of contracting this disease.

Jacqui A. Shykoff, Erika Bucheli

Oliver Kaltz

*Swiss Federal Institute of Technology,
Experimental Ecology, ETHZ - NW,
8092 Zürich, Switzerland*

Flower lifespan and disease risk

SIR—Ashman and Schoen^{1,2} have proposed that plant fitness is an increasing function of pollination success, and a decreasing function of floral maintenance and production costs. They concluded that different pollination regimes, and different relative costs of flower construction and maintenance, determine optimal floral longevity for a species or population¹. Here we suggest another factor relevant to the evolution of flower lifespans.

The pollinator-borne plant pathogen, the anther-smut *Microbotryum violaceum* (Pers.) Deml and Oberw, affects members of the pink family (Caryophyllaceae³) and is mainly transmitted by pollinators⁴. Successful infection requires several time-consuming processes. After a diploid teliospore, the dispersal unit of this basidiomycete fungus is carried to an uninfected plant by a pollinator, it undergoes meiosis, and two haploid sporidia of opposite mating type conjugate to produce the infectious dikaryon that invades the plant⁵.

Flowers that remain open or on the plant longer are at higher risk of contracting this disease than are shorter-lived flowers, for at least two reasons. Longer-lived flowers have an increased chance of receiving infectious spores due to increased numbers of insect visits, as shown by gradual pollen accrual over time⁶, and longer-lived flowers offer the fungus a greater opportunity to colonize the plant. This is true both of flowers that are maintained open and those that have wilted but retain physiological connections with the plant. Ripening fruits of female and perfect flowers remain physiologically dependent on the plant. Male plants, on the other hand, can shed their flowers completely, losing nothing once their pollen is removed.

We measured floral longevity in full sib crosses of two dioecious species of the family Caryophyllaceae—the white and red campion, *Silene latifolia* and *S. dioica*. Both species harbour this disease, which may have been instrumental in the evolution of floral traits⁷. Male flowers remained on the plant for one or two days only (*S. latifolia*, 1.5 ± 0.05 days; *S. dioica*, 1.1 ± 0.05 days; mean $\pm$ s.e.); female flowers wilted soon

after ovule fertilization, but retained their physiological connections for a much longer time (*S. latifolia*, 8.9 ± 0.27 days; *S. dioica*, 9.6 ± 0.40 days).

Female plants seem to be more prone to anther-smut in natural populations. Field surveys throughout Switzerland showed that females of these dioecious species had significantly higher infection rates than did males. In six naturally infected populations of *S. latifolia*, on average 7.4% of females, but only 4.2% of males, were infected (analysis of deviance, $F_{(1,4)} = 36.03$; $P < 0.005$). Similarly, in three populations of *S. dioica*, on average 3.6% of the female plants were infected compared with only 1% of the males ($F_{(1,2)} = 24.88$; $P < 0.05$). Higher infection of females in natural populations may have other explanations, such as higher mortality rates of infected male than infected female plants (A. Biere, personal communication), but females may be more susceptible to this fungal pathogen, simply because their flowers remain longer on the plants, while it may have been a selective force shaping the rapid phenology of male flowers.

A logical extension of this hypothesis is that populations or species with a higher risk of infection should have shorter optimal floral lifespans. Infection is known almost exclusively from perennial members of the Caryophyllaceae³, so we predict that annual species in this family should have longer floral longevity, as will any populations that have evolved in the absence of disease. We also suggest that anther-smut disease may facilitate the

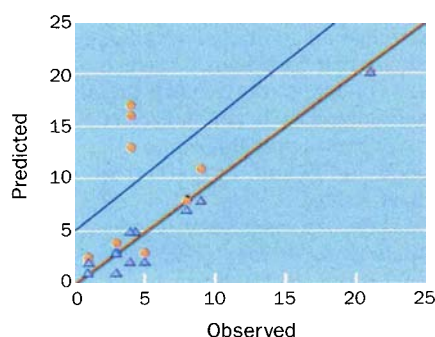
SIR—A recent article¹ proposed a model of flower longevity based on the number of flowers per plant, the cost to produce and maintain a flower, and the male and female fitness accrual rates. I contend that potential problems in the authors' field study of 11 plant species have been underestimated. First, data were collected only during one season and at one location, and so may not adequately represent the typical responses of the species studied. Second, the longevity data were obtained without controlling pollinator access. I would challenge the proposition that floral longevity measured in the field is a good measure of maximum floral longevity in the absence of pollinators. Many flowers become unattractive to pollinators shortly after pollination^{8,9}, so, in the absence of any control on pollinator access, flower longevity may simply be a measure of how long it took the flower to become pollinated.

The percentage variance in reported data accounted for by the model (r^2) was only 0.40 (see figure), but when the flower longevity data are plotted against the time to 80% emptying of the anthers (from Fig. 2 of ref. 1), a much better fit is found ($r^2 = 0.94$; see figure). The authors appear to have determined, at least partly, pollinator activity in flowers whose life is pollination-dependent.

The effect of pollinators on floral longevity is mediated by a complex internal signalling system, in which ethene (ethylene) is a key molecule^{9,10}. Exogenous ethene usually mimics the effects of pollination on flower petals^{10,11}. Regulation of flower life by ethene has been found to be similar within plant families; for example, ethene can hasten perianth abscission or wilting¹⁰. Some families, by contrast, show flower wilting that is independent of ethene¹⁰, and no examples are apparently known of flower life shortened by pollina-

1. Ashman, T.-L. & Schoen, D. J. *Nature* **371**, 788–791 (1994).
2. Schoen, D. J. & Ashman, T.-L. *Evolution* **49**, 131–139 (1995).
3. Thrall, P. H., Biere, A. & Antonovics, J. *J. Ecol.* **81**, 489–498 (1993).
4. Roche, B. M., Alexander, H. M. & Maltby, A. D. *Ecology* **76**, 1863–1871 (1995).
5. Day, A. W. & Garber, E. D. in *Genetics of Plant Pathogenic Fungi* (ed. Sidhu, G. S.) 457–482 (Academic, London, 1988).
6. Shykoff, J. A. *Am. J. Bot.* **79**, 138–144 (1992).
7. Elmquist, T., Liu, D., Carlsson, U. & Giles, B. *Oikos* **68**, 207–216 (1993).
8. Gori, D. F. in *Handbook of Experimental Pollination Biology* (eds Jones, C. E. & Little, R. J.) 31–49 (Van Nostrand Reinhold, New York, 1983).

9. Stead, A. D. & van Doorn, W. G. in *Molecular and Cellular Aspects of Plant Reproduction* (eds Scott, R. J. & Stead, A. D.) 215–237 (Cambridge Univ. Press, UK, 1994).
10. Woltering, E. J. & van Doorn, W. G. *J. exp. Bot.* **39**, 1605–1616 (1988).
11. Stead, A. D. *Pl. Growth Regul.* **11**, 13–20 (1992).
12. Nichols, R. J. *hort. Sci.* **41**, 279–290 (1966).
13. Wacker, H. *Jb. wiss. Bot.* **49**, 522–578 (1911).
14. Fitting, H. *Naturwissenschaften* **9**, 1–9 (1921).
15. Eisikowitch, D. & Lazar, Z. *Bot. J. Linn. Soc.* **95**, 101–111 (1987).
16. Kerner von Marilaun, A. *Pflanzenleben* vol. 2 (Bibliograph. Inst., Leipzig, 1891).
17. Ashman, T.-L. & Schoen, D. J. in *Floral Biology* (eds Lloyd, D. G. & Barrett, S. C. H.) 112–139 (Chapman & Hall, New York, 1996).



Relationship between predicted and observed flower longevity in 11 species. The red line is the ideal relationship. Longevity as observed in the field is plotted against predicted values; linear regression is shown as a blue or green line. Values predicted based on ref. 1 (circles, $r^2 = 0.40$; blue line), or based on 80% emptying of the anthers (triangles, $r^2 = 0.94$; green line).

tion, in these families (for example, ref. 11). Ethene, furthermore, causes the loss of other cues for pollinators, such as colour¹¹ and permanent flower closure¹². It is likely that the principal function of the regulation of flower longevity by ethene is to direct pollinator activity; avoidance of pollinated flowers increases the time spent by pollinators on virgin flowers⁸.

Petal wilting or abscission strongly reduce visual 'advertising' by flowers, and are usually considered to conclude floral life. Permanent flower closure often precedes wilting or abscission, and also reduces the visual presence of flowers. It is unclear whether the authors also considered permanent closure to end flower life, but in the present context, I define floral life as the time from (first) opening to petal wilting or abscission.

Among the species investigated by Ashman and Schoen¹, at least five belong to families in which floral longevity is shortened by ethene¹⁰; one is from the Boraginaceae, a family known to show petal abscission¹³ (hence their flower life seems ethene-regulated) and in which the flower life of *Borago* sp. is shortened by pollination¹⁴, and another, *Oenothera flava*, is from a genus known to contain at least one species whose flower life is shorter after pollination¹⁵.

The male and female fitness accrual rates, as used by the authors, are also a function of pollinator activity, as male fitness accrual rate was defined as "the rate at which pollen is disseminated and enters the pool of pollen that competes to fertilize ovules", and female fitness accrual as "the rate at which pollen is received to fertilize ovules".

More than a century ago, Kerner von Marilaun¹⁶ suggested that flower life was related to plant and flower morphology, and to the chance of reproductive success. He noted that flowers with many anthers and much pollen tend to have a short life, whereas those with one anther are long-lived, particularly when their pollen is all

in one package. Flower life in species with many flowers per individual also tends to be short, especially when these open sequentially over a long period. In contrast, flowers are relatively long-lived on plants that produce only one or a few blooms per individual. A long flower life was also predicted for species that experience few pollinator visits due to pollinator specialization, and for species with obligate outcrossing. These hypotheses have, however, never been properly tested.

An insight into these relationships will require flower longevity to be monitored at several sites and over several seasons. In addition, cessation of pollinator cues, such as colour (including the ultraviolet range) and permanent flower closure, along with wilting or abscission of the perianth, should be taken into consideration. In these studies, access by pollinators must be regulated to enable intrinsic flower longevity to be distinguished from the shorter flower life that may occur in the presence of pollinators.

Wouter G. van Doorn

Agrotechnological Research

Institute (ATO-DLO), PO Box 17,

6700 AA Wageningen, The Netherlands

ASHMAN AND SCHOEN REPLY — Shykoff *et al.* suggest that because vulnerability to receipt of the anther-smut pathogen (as well as subsequent infection) increases as flowers remain open and physiologically connected to the plant, natural selection should operate to reduce floral longevity, and should do so more in males than in females, where longer-duration floral connections are required for completion of fertilization. We agree that these plant-parasite interactions could influence floral longevity in the species of *Silene* studied. The essential elements of interaction could, in fact, be easily accommodated by extending our optimal floral longevity model¹ to consider the rate at which the infectious spores arrive at flowers during the pollination process, and the extent to which disease infection initiated at one point in time diminishes future reproductive success. Moreover, because male and female reproductive functions are separated in dioecious species, there would be no constraint on male floral longevity (in the case of plants with bisexual flowers) brought about through possibly slower rates of fitness accrual of female fitness. Thus, factors that select for reduced floral longevity, such as increasing disease susceptibility with increasing floral lifespan, would be free to lead to divergence in optimal male and female flower longevities.

With regard to the comment by van Doorn, we have acknowledged elsewhere that post-pollination responses, such as ethylene production and floral-senescence response, can contribute to variation in floral longevity among and within species¹⁷. There is, however, a close corre-

lation between floral longevity measured in flowers protected from pollination (MFL) and floral longevity in flowers of the same species that have been exposed to pollinators (RFL) ($MFL = 0.88(RFL) + 1.72$; $P < 0.0001$; $n = 23$ plant species), suggesting that using either measure of longevity would give results that are compatible with our model's predictions.

van Doorn's suggestion that variation in floral longevity can be accounted for primarily by floral senescence in response to ethylene production is at odds with the fact that flowers of different species differ in durability (and therefore longevity). For example, the flowers of morning glory (and other species), whether pollinated or not, senesce on the same day they are produced. It has been suggested that 'pollination-induced' senescence is mediated by production of ethylene by the stigma¹⁰. If such a mechanism is the main determinant of floral longevity, we would expect a strict correlation between completion of female function and floral longevity. van Doorn re-analyses our data and reports that the correlation between observed longevity and "time to 80% emptying of the anthers" is better than with floral longevities predicted by our model. However, when one uses the more appropriate measure of completion of female function, which we provided in our original data (that is, time to 80% receipt of pollen required for full seed set), the correlation is far less supportive of his hypothesis ($r^2 = 0.16$; $P > 0.2$).

We agree that it would be valuable to collect more data on floral longevity and its relationship to male and female fitness accrual from additional populations and species. We have, in fact, analysed floral longevity and correlated male and female fitness accrual in additional species, and the results support the predictions of our model⁴.

Finally, there is a deeper issue. van Doorn argues for the primacy of a proximate explanation of floral longevity, rather than for an ultimate (evolutionary) explanation. This view that floral longevity variation is simply a physiological problem misses the issue of why plants have evolved mechanisms to control floral longevity. We believe that this ability came about from past selection to optimize floral longevity in the face of varying floral maintenance costs and varying rates of contributions of flowers to fitness. It would not be surprising or at odds with our model that cues such as ethylene production could be used by plants to realize such optimal floral longevities.

Tia-Lynn Ashman

Department of Biological Sciences,

University of Pittsburgh,

Pittsburgh, Pennsylvania 15260, USA

Daniel Schoen

Department of Biology,

McGill University,

Montreal, Quebec H3A 1B1, Canada

From durable to disposable

Howard P. Segal

American Plastic: A Cultural History. By Jeffrey L. Meikle. Rutgers University Press: 1996. Pp. 403. \$49.95.

In the opening scene of the 1967 film "The Graduate", Dustin Hoffman receives career advice from a family friend: "I just want to say one word to you. Just one word... Plastics... There's a great future in plastics." In the context of the film, 'plastics' represents not just the excessive materialism of the age that Hoffman's character rebels against but also the superficiality of plastics as a product and the illusion that plastics could be taken seriously as an occupational calling. Not surprisingly, the plastics industry has never quite recovered from that memorable putdown. Yet people born after the movie appeared, Jeffrey Meikle observes, generally have no sense of plastics as both a pejorative word and a questionable product.

Meikle explains this generational gap between the 1960s and the 1990s while offering a landmark account of a phenomenon that dates back to the late nineteenth century and continues through to the present. He combines a first-rate technological history with a most impressive cultural analysis of how plastics evolved from a material surrounded by utopian expectations to a material epitomizing inferiority and eventually to a part of everyday life barely noticed except when environmental concerns and crises bring it public attention.

Meikle details the history of plastics from the invention of celluloid in 1869 as a substitute for ivory in billiard balls, through the introduction of synthetic materials such as Bakelite, Plexiglas, vinyl and polystyrene in the early and mid-twentieth century, to the replacement of the 'Plastic Age' by the 'Information Age' in recent decades. Derived from vast research, his lucid and illuminating discussions would themselves suffice for many historians of technology. But Meikle has greater ambitions. He seeks to understand why so many Americans have long been ambivalent about plastic as a cultural symbol, why such plastic products as Bakelite radios and nylon stockings have always been widely appreciated, while others, for example Tupperware

and polyester suits, have eventually evoked ridicule. At the deepest level, Meikle investigates why contemporary culture and identity are themselves increasingly perceived as plastic, as ever "more malleable, less permanent, even ephemeral". In the process, he considers the relationship between that perception



Quality furniture leads a charmed and charming life surfaced with
Catalin LAMINATING RESINS

Out to lunch: an all-plastic dining room, 1963.

and the immaterial cyberspace of today's electronic media.

Not content with limiting his research to actual objects and technical and industrial publications, Meikle has included representations of plastics in art, literature, film and other areas of culture. Nor has he ignored philosophy: among his most interesting insights is the degree to which the "plasticity of experience" heralded by such quintessentially American philosophers as Ralph Waldo Emerson and William James has by now been reinforced by the incorporation of plastics in our common consciousness.

A professor of American studies and art history at the University of Texas at Austin, Meikle demonstrated his gift for making the history of technology and culture accessible to the proverbial general

reader in his first book, *Twentieth Century Limited: Industrial Design in America, 1925-1939*. There are, in fact, connections between that work and *American Plastic*. For one thing, industrial design became a recognized profession at the same time — the Great Depression — that plastics became an identifiable industry. Such prominent industrial designers as Henry Dreyfuss, Norman Bel Geddes, Raymond Loewy and Walter Dorwin Teague used plastics in their various projects. Like these designers, the foremost advocates of plastics believed that, despite economic disaster and social despair, they could nevertheless manufacture a veritably perfect world that would somehow keep getting better — and so keep the industry flourishing. They failed to realize that the achievement of their goal, Utopia, the epitome of stasis, would put them out of business. Neither the industrial design profession nor the plastics industry ultimately confronted this dilemma once popular tastes changed, as they invariably did, contrary to the naive expectations of both groups that they would remain permanently avant-garde. But Meikle is sensitive to subtle shifts over time within plastics' utopian prophecies. As earlier dreams of ultimate social stability gave way to later dreams of unceasing social change, plastics were less often retained and revered and more frequently disposed of and readily replaced.

Meikle is equally sensitive to the role of plastics in asserting our unprecedented control over nature, first as a triumph of civilization, more recently as at best a mixed blessing. The growing sense in the United States and elsewhere that we live in a ruined

world of our own making can hardly be blamed on plastics alone, but plastics as environmentally harmful wastes have certainly become a focal point for such sentiments.

Like *The Real Thing: Imitation and Authenticity in American Culture, 1880-1940* by Miles Orvell, *American Plastic* takes up, albeit more peripherally, the profound changes that took place as American popular culture shifted from an emphasis on imitation, on the newly available mechanical reproduction of familiar objects and experiences for the masses (furniture, architecture, artworks, photographs and so on) to an emphasis on authenticity, on the use of similar technologies to create new objects and experiences. Yet Meikle, like Orvell, recognizes that the culture of authenticity never

wholly replaced that of imitation. If anything, the passion for imitation persists not only in the many objects Meikle discusses but also in enterprises such as the Disney and kindred pseudo-historical theme parks he also examines.

American Plastic could be faulted for insufficient comparisons with developments outside the United States, an all too common failing of otherwise outstanding works in American studies like this one. And the publisher could be faulted for the quality of the book's many and otherwise very useful black-and-white illustrations, although its several pages of handsome colour plates are a partial compensation.

But these disappointments pale beside what Meikle has provided: one of the most significant works ever written in the history of American technology and culture. □

Howard P. Segal is in the Department of History, University of Maine, Orono, Maine 04469-5774, USA.

It's no go Blavatsky

Walter Gratzner

An Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural.
By James Randi. *St Martin's*: 1995.
Pp. 284. \$24.95.

A DECADE or two ago there appeared in a newspaper an interview with an aged but sprightly inventor. The old fellow reminisced that as a young man during the First World War he had put before the Admiralty a sure-fire scheme for eliminating the shipping losses in the North Atlantic, which were then imperilling the nation. He had conceived of a gigantic hovering platform, which having been loaded in the United States with war materials would be made to rise some feet above the Earth's surface. It would then be merely a matter of waiting for the Earth to rotate and England to appear below. The inventor had been chagrined by the cool reception that this ingenious proposition received at the time, but he understood now that the device was based on a false premise, namely that the Earth was a sphere. For he had recently succeeded to the presidency of the Flat Earthers. Since then, Britain's leadership has, as in so many respects, been surrendered to the United States: you will find the Flat Earth Society's address in California recorded in James Randi's book. The Hollow Earth theory too, which tells us, still has its adherents, especially in Germany. It was conceived by Edmund Halley and refined by a succes-

sion of savants over the centuries, and in its most advanced form teaches that we live on the inside of the spherical shell with the Sun at its centre, although we see only its reflection from the walls. There are openings at the poles, but nothing at all outside.

James ('the Amazing') Randi is distinguished as a magician and writer, the nemesis of spoon-benders, psychics and telekineticists, and a tireless opponent of

IMAGE
UNIMAGE
FOR COPYRIGHT
REASONS
REASONS

Upwardly mobile: artist's impression of a reported 'alien abduction' event in 1989.

credulity and unreason. His excellent compendium is arranged alphabetically, copiously cross-referenced ("aetite — see bezoar", a stone found in the entrails of animals, possessing miraculous curative properties) and enlivened by many pungent asides: 'conjureting', we learn, is the ability to call up demons and storms — not, Randi opines, a socially admired talent.

After a year that saw the largest British building society (to Americans, a thrift) issue horoscope-based guidelines for investors, Orange County, California, bankrupted by a treasurer who gambled on futures, secure in the advice of an astrologer, and a statue of the Virgin weeping tears of blood, who can argue that Randi's efforts are not needed? Nor does exposure dismay the believers: the term "mixed mediumship" has even been coined to describe mediums who have been caught cheating — when the deception has not been spotted, the act, is of course counted as genuine. Truth is not an easily marketable commodity.

Such hare-brained gullibility is by no means unknown among members of the

learned professions; there are professors of parapsychology in ancient seats of learning and they have validated the claims of all manner of impostors from mind-readers to benders of cutlery. Randi gives details of two two psychics, Kulagina and Kulashova by name, who practised blindfold reading tricks (the currency of magicians the world over) and towed matchboxes on thin nylon threads to such effect that in 1978 the Soviet Academy of

Michael Bühler/MEPL

Sciences upheld their claims to paranormal powers. Psychiatrists have re-invented Satanic abuse of children and "hidden memory" and have been enacting reruns of the Salem witch trials. And at Harvard in 1993, Dr John Mack published the results of his investigations into cases of abduction by extraterrestrial aliens: his subjects, he concluded, had indeed probably been borne aloft in UFOs for purposes of procreation. Dr Mack's opinions on the Tooth Fairy, Randi adds, are not yet revealed.

Randi has unearthed many strange historical episodes. Totalitarian regimes have always been easy but dangerous prey to charlatans and monomaniacs. The celebrated medium Madame Blavatsky and her racial categories, which included Aryans as the 'Race of Hope', had a profound influence on Himmler. Two clairvoyants rose to positions of prominence in Hitler's Germany: Karl Krafft and Erik Jan Hanussen (*né* Herschel Steinschneider) predicted brilliant successes for the Nazi party and for German arms and were patronized by Hitler, Himmler, Goebbels and Hess. Both overreached themselves, one by predicting all too accurately the Reichstag fire, the other the Munich beer-hall attempt on Hitler's life. The Gestapo must have harboured sceptics among its thugs, for both soothsayers were arrested and met nasty ends.

The most unsettling message to emerge from Randi's researches is that superstition and ignorance are on the increase. Bizarre cults flourish in lush profusion. Lecherous and venal tele-evangelists mulct their followers in unprecedented numbers, aromatherapy and colonic irrigation clinics thrive and jets daily convey their loads of pilgrims to Lourdes. Much of what Randi reveals is of course high comedy. He has missed a few opportunities: scientology, one of the funniest of all cults, receives a disappointingly brief mention; Arthur Koestler is remembered, but not the ineffable Sir Alastair Hardy; the Urantia cult, about which Martin Gardner has written a hefty book, does not feature;

and surely Rupert Sheldrake has earned a mention. To compensate there are two appendices, one of which reveals that the Curse of King Tutankhamun's Tomb can take the form only of increasing the longevity of trespassers, the other enumerating the dates predicted over the centuries for the End of the World. (The next one comes up this year and another in 1999). Randi stops just short of including religions generally and their versions of the deity among his hoaxes and delusions, although their paraphernalia — saints and miracles, angels, devils and suchlike — abound. I suspect he would (as do I) concur with Lord Eldon, the Lord Chancellor who, having ploughed through *Paradise Lost*, gave his opinion of Satan thus: "Damn fine fellow, and I hope he wins". □

Walter Gratzer is at the MRC Muscle and Cell Motility Unit, King's College, 26–29 Drury Lane, London WC2B 5RL, UK. He is editor of *A Bedside Nature: Genius and Eccentricity in Science 1869–1953*, an anthology of extracts from Nature recently published by Macmillan. Price is £19.95, \$29.95 (nonsubscribers); £14.95, \$24.50 (subscribers). Further details available from Nature on +44(0)171 843 4962 (tel.).

Techniques of extraction

Robert W. Cahn

Stalin's Captive: Nikolaus Riehl and the Soviet Race for the Bomb. By Nikolaus Riehl and Frederick Seitz. *American Chemical Society*: 1995. Pp. 218. \$34.95.

NIKOLAUS Riehl was a German radio-chemist, born in St Petersburg in 1901, who in 1927 obtained his doctorate in Otto Hahn's laboratory under Lise Meitner's supervision. He then joined the Auer Company (founded by Auer von Welsbach, the originator of the eponymous gas mantle) became its research director and before the Second World War had invented the fluorescent electric lamp. The Auer Company worked with uranium ores, from which it extracted radium and rare-earth oxides, so Riehl had the chemical expertise needed to purify metallic uranium when the war came. He became fascinated by the idea of a chain reaction, putting his services at the disposal of the German 'Uranium Club' (or nuclear project). At the end of the war, the Russians (specifically Beria's secret police) captured Riehl and his family and took them to Russia so he could teach engineers there how to purify uranium. He was forced to remain in the country until 1955, when he returned with his

THE migrant hawk dragonfly *Aeshna mixta* is common throughout most of south and central Europe, and is also found in North Africa, the Caucasus and east to China and Japan. Britain and Ireland support about a third of the European dragonfly fauna. The status of all their resident and more frequent immigrant species over the past 25 years is now surveyed in *Atlas of the Dragonflies of Britain and Ireland* by R. Merritt, N. W. Moore and B. C. Eversham. The atlas provides detailed information about distribution, descriptions to aid identification and summaries of dragonfly behaviour, habitat and conservation. HMSO, £15.95 (pbk).



family to West Germany. Soon after, he wrote an account (in German) of his experiences, *Ten Years in a Golden Cage*, which was published in 1988 but made no impact at all. He died in 1990.

The book came to the attention of Frederick Seitz, an influential US physicist (it would be no exaggeration to consider him the father of solid-state physics). Having belonged to the pre-war generation of physics students who, as he points out, were required to learn German to a good standard of competence, Seitz was able to translate the book himself and he persuaded the American Chemical Society to publish it as an historical treatise. The translated memoirs are preceded by an extraordinarily wide-ranging 66-page preamble written by Seitz, which places Riehl in perspective in relation to the early work on fission, the Uranium Club and several other eminent scientists such as the physical chemist Max Volmer (of 'nucleation' fame) and Gustav Hertz, both of whom were also taken to Russia and whose lives crossed with Riehl's. Several pages of the preamble deal with the much debated issues of Heisenberg's role in the German nuclear effort and the Farm Hall transcripts, the surreptitiously recorded conversations of the ten German physicists detained in England from May to December 1945. This preamble is so full of fascinating nuggets of information, scientific and political, that it would be well worth buying the book just for that.

Riehl was fluent in German and Russian (his mother was Russian) and therefore understood much about his Soviet experience that would have escaped a monoglot. His memoirs paint a

remarkable human picture, both of ordinary people (and guards) in his immediate environment and of the officials, engineers and managers with whom he had to interact professionally. For instance, there is an intriguing passage in which a senior engineer pushes, with unaccountable resolve, a novel chemical purification technique of which Riehl had never heard, and from the engineer's use of words he divines that he is receiving the clandestine results of Soviet espionage in the United States. Indeed, there is much about Russian vocabulary, including the systematic use of foul language to achieve political ends. He is also most illuminating on the pointless extremes to which the Soviet authorities pushed their obsession with secrecy: it is a tribute to Riehl's subtle understanding of Russian psychology that he could finally persuade them to let him go (perhaps the fact that his expertise was trumped by stolen American know-how also accelerated his liberation). We also receive vivid pen portraits of the different surroundings of his varied abodes, from near Moscow to the Crimean coast.

Stalin's Captive can be comfortably read in an afternoon, and its intrinsic digestibility is further enhanced by the many portrait photographs of a range of eminent scientists (Seitz is a past master at locating these). The book is recommended to all sorts and conditions of readers; no-one could find it boring. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.

Betrayers of trust?

John Galloway

The Power of Life or Death: A Critique of Medical Tyranny. By Fabian Tassano. Duckworth: 1995. Pp. 177. £12.95 (pbk).

Wrong Medicine: Doctors, Patients and Futile Treatment. By Lawrence I. Schneiderman and Nancy S. Jecker. Johns Hopkins University Press: 1995. Pp. 200. \$25.95, £21.50.

THE trouble with medicine is that it has shifted from magic and mystery to science and technology without society, including the medical profession, quite realizing this fact. Science in the guise of medical technology now gives medicine considerable power over people's lives, including their deaths. Yet it still seems that medical training involves an initiation into arcane mysteries rather than learning practical and technical skills and judging if and when to use them. Society accords doctors a considerable moral authority not necessarily granted to other professionals. The powerful concept of clinical freedom has its roots here with its implication that doctors are never wrong — at worst only negligent.

When medicine was largely ineffectual, this hardly mattered. The trouble started when science began to make medicine work. Serious moral and economic issues arise when one really can prolong or 'save' someone's life rather than letting nature take its course. At this point, matters of 'freedom' and 'choice' shift to a real agenda. Because this agenda includes the question of resources, which never seem sufficient to meet the demand or need for them, it raises the interesting conflict between collective and individual rights and the implications for doctors' ethical codes. Who should make decisions about life or death and on what grounds? These two books attempt to deal with these questions.

Fabian Tassano, an economist based in the United Kingdom, is concerned that doctors make decisions about life and death that they have no right to be making. He condemns both medicine's monopoly and its paternalism, as typified by the attitude of "I'm a doctor and I know best". He argues that statutory self-regulation of the medical profession, which ostensibly protects the public, in fact protects the monopoly.

The author shows clearly how society buttresses the view of doctors as moral arbiters and the gate-keepers of health care — and so of health; there is, for instance, a marked tendency for courts to favour doctors' views rather than those of patients and their families when differences of opinion arise over a decision to

prolong or end life. Indeed, Tassano seems to be a strong follower of George Bernard Shaw, who believed that "All professions are a conspiracy against the laity".

At the core of the book, however, is Tassano's support of the rights of the individual. He believes that doctors lean towards a collectivist morality that puts us all individually at risk. He argues strongly that doctors owe a duty only to individual patients and that concerns about the public interest are at best irrelevant and at worst downright dangerous. He thinks that notions of equity are misplaced. Because equity of access to care is at the forefront of the UK National Health Service, he has no truck with it. If a patient wants treatment and can pay for it, then Tassano sees no reason why the patient should not get it.

His view goes straight to the medical jugular. Is medicine after all just a commodity, like baked beans or cars, subject only to the laws, morals and ethical codes of the marketplace? Are doctors' claims to be 'special' just a good marketing ploy? Undeniably there is a strong profit motive in medicine, and rackets are certainly in evidence, as alluded to in the book by Lawrence Schneiderman, a doctor, and Nancy Jecker, an ethicist. One feature of medicine in the marketplace raised by these US authors is that, in common with other markets, what is on offer may be worthless. The old habit of not questioning a form of treatment may give rise to the illusion that the treatment works. That science makes for effective medicine is sometimes true, but not always. Even when it works, medical science may create more problems than it solves. Medical science is slow to recognize what is demonstrably wrong or to accept what has been proved right. Why do doctors continue to give futile treatments? Do they find it impossible to admit they are powerless to help? Is the treatment simply symbolic as the authors suggest? And do patients collude with doctors in their expectations of 'a pill for every ill'? (Science itself is hardly blameless here, with its continuous stream of well-publicized 'breakthroughs').

Extreme cases reveal underlying principles most clearly. Both books are littered with examples of the moral, emotional and financial thickets ensnaring life-and-death decisions. Should medicine help you if the treatment will cause you intense suffering and is unlikely to cure you? Should the choice be yours? Should a baby with spina bifida be helped to live briefly, be allowed to die or even be killed? If the baby is likely to die very soon in considerable pain, would it not be better for that time to be as short as possible and for the death to be controlled? How does the parents' suffering enter the equation? How important are their views? If you wish to end your life, should

a doctor help you? Should someone in a coma with irreparable brain damage be kept physiologically alive and, if so, for how long?

Science does not come with instructions on use; it has forced medicine and society to face the fact that decisions about life and death have to be made, but says nothing about who should make them. Surely this is a question for society and its individual members. Doctors and patients must all be well informed about care and its consequences and their views must be based on the best possible information. We seem to have got as far as the idea of evidence-based medicine. Let's go all the way and have evidence-based patient choice. □

John Galloway is at the Eastman Dental Institute and Hospital for Oral Health Sciences, 256 Gray's Inn Road, London WC1X 8LD, UK.

Dictionary digest

Latin Names Explained: A Guide to the Scientific Classification of Reptiles, Birds and Mammals by A. F. Gotch.

Cassell, £20. "A mine of information and illumination to naturalists who have used these names for years without understanding their import", wrote *Nature's* reviewer of an earlier edition.

Ainsworth and Bisby's Dictionary of the Fungi by D. L. Hawksworth, P. M. Kirk, B. C. Sutton and D. N. Pegler (8th edn).

CAB International, \$49.95, £30. Contains some 20,000 entries, covering all fungal generic names and terms, biographical notes, information on metabolites and mycotoxins and broad accounts of pure and applied aspects of the subject. Rich pickings for professional and amateur mycologists alike.

The Concise Oxford Dictionary of Ecology edited by Michael Allaby. Oxford University Press, £7.99 (pbk). Contains 5,000 entries derived from the *Oxford Dictionary of Natural History*. Includes biographical notes on important figures in the subject.

The Concise Oxford Dictionary of Earth Sciences edited by Ailsa Allaby and Michael Allaby. OUP, £8.99 (pbk).

Nature's reviewer marvelled at "how the editors have compressed so much so clearly... a feat which occurs only once in a blue moon — and even that is defined!".

Correction

In Irving M. Klotz's review of *Nazi Science* by Mark Walker (*Nature* 379, 410; 1996), the Von Weizsacker referred to should have been Carl Friedrich, not Ernst (his father and Nazi Minister of State). The mistake was made in the editorial office of *Nature*. Our apologies.

Geophysical evidence for lithospheric delamination beneath the Alboran Sea and Rif–Betic mountains

Dogan Seber*, Muawia Barazangi*, Aomar Ibenbrahim† & Ahmed Demnati‡

* Department of Geological Sciences and Institute for the Study of the Continents, Cornell University, Snee Hall, Ithaca, New York 14853, USA

† Centre National de Coordination et de Planification de la Recherche Scientifique et Technique, BP 1346, Rabat, Morocco

‡ Office National De Recherches Et D'Exploitations Pétrolières, BP 8030, Rabat, Morocco

Geophysical evidence is presented for an episode of active delamination of a piece of continental lithosphere. Observations of earthquake hypocentre locations, seismic wave velocities and attenuation, Bouguer gravity, seismic reflection and drill hole data are combined with surface geology to infer the presence of a high-velocity, seismically active, rigid body in the upper mantle beneath the Alboran Sea and surrounding Betic and Rif mountain belts of the western Mediterranean region. This upper-mantle body, inferred to be the delaminating continental lithosphere, is overlain by a low-velocity, aseismic and strongly attenuating uppermost mantle, inferred to be the asthenospheric material replacing the delaminating lithosphere.

THE concept of lithospheric delamination and its place in the geodynamic evolution history of the Earth have recently received increased attention among geoscientists. It was proposed that extensional collapse in old orogenic belts is a consequence of removal of lithospheric material^{1,2}. Two separate models were put forward to explain this extensional collapse: a delamination model³ and a convective downwelling model⁴. In the former model, the subcrustal lithosphere detaches from the overlying crust and peels away exposing subcrustal depths to hot asthenospheric material. In the latter model, the thickened lithosphere in continental collision zones detaches and sinks into the asthenosphere (possibly as a whole) owing to gravitational instability created by thickening (see Fig. 1 of ref. 5). The term “delamination” is used loosely in the literature to refer to removal of lithospheric material by either of these processes. Here we also use delamination to refer to a physical phenomenon that detaches pieces of lithosphere by one of these processes. But when we refer to delamination as in the former model above, we state it as “delamination (*sensu stricto*, *s.s.*)”. Both of these proposed models predict a sudden uplift at the surface following the detachment of the lithosphere and subsequent extensional collapse and extensive volcanism⁶. Geological observations of this type have been used to propose potential sites for lithospheric delamination (for example, the Basin-and-Range Province and the Colorado plateau, the Appalachians, the Pannonian basin, the Tibetan plateau, the Puna, and the Alboran Sea). To date, however, no direct geophysical observations (for example, seismicity, velocities or attenuation), which are expected consequences of delamination, have been presented in a comprehensive manner from any of these regions.

Based on our geophysical data and analyses presented here along with previous geological interpretations, we provide sound evidence for lithospheric delamination beneath the Alboran Sea and the surrounding Rif and Betic mountain belts. The enigmatic occurrence of intermediate-depth earthquakes and their spatial distribution mapped in this study are among the most significant geophysical evidence for active delamination. A gap in seismic activity between depths of about 20 and 60 km beneath the Alboran Sea coincides with a region of low P_n velocities and

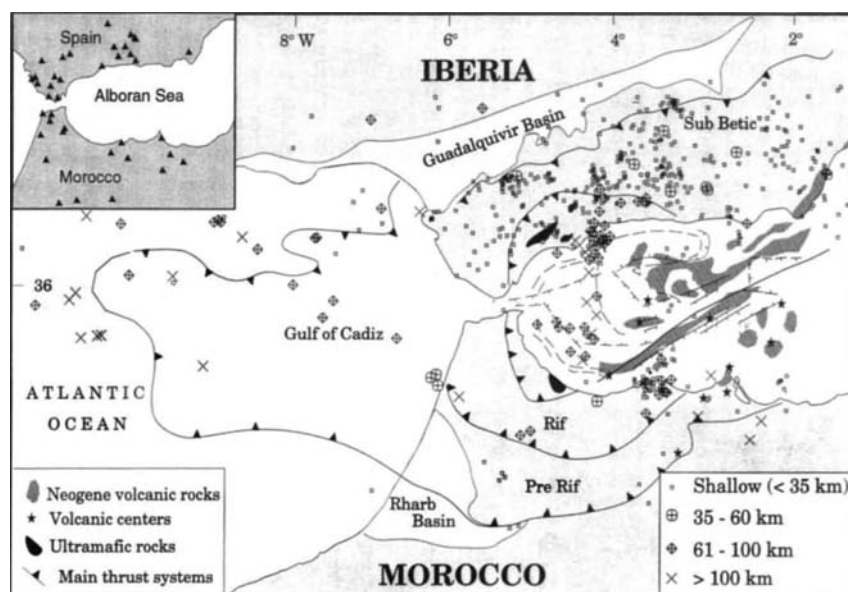
high seismic wave attenuation (see Fig. 3 legend for seismic wave nomenclature). Teleseismic P-wave residuals and tomographic images indicate a high-velocity upper-mantle material beneath this low-velocity region^{7,8}. The presence of the high-velocity body, which is seismically active, beneath the anomalously low-velocity, high-attenuation zone is used to argue that asthenospheric material is underlain by a rigid upper mantle which we interpret to be the delaminated continental lithosphere. Pronounced low Bouguer gravity anomalies in the region and their three-dimensional modelling, constrained by drill-hole and seismic-reflection data, also support the presence of low-density material in the lower crust and/or uppermost mantle levels. High heat-flow data and volcanism in the region are also in agreement with this proposed model.

Both surface and sub-surface structural symmetry along an east–west line across the Strait of Gibraltar is in support of the convective removal model, whereas spatial distribution of subcrustal earthquakes and seismic wave attenuation in the uppermost mantle is more suggestive of the delamination (*s.s.*) model. The truth probably lies between these two models. A convective removal may have started the phenomenon, and the instability within the lithosphere is now a driving mechanism for lithospheric delamination (*s.s.*). It should be acknowledged that the delamination (*s.s.*) in this case has a three-dimensional geometry; that is, Iberian lithosphere is peeling away to the north, while the Moroccan lithosphere is peeling away to the south, and the Gulf of Cadiz lithosphere is peeling away towards the west into the Atlantic. Beneath the eastern Alboran Sea region we observe no evidence of lithospheric mantle that is attached to the overlying crust. We interpret this structure to indicate that the mantle lithosphere in this region was removed at an earlier stage.

Geological setting

The Rif mountain belt of northern Morocco and the Betic belt of southern Spain (that is, the Gibraltar arc) are mountain belts of Alpine orogeny that surround the Alboran Sea basin of the western Mediterranean (Fig. 1). These thrust belts are characterized by south-, west- and north-vergent low-angle thrust systems with tectonic transportation directed away from the basin^{9–11}. The

FIG. 1 Generalized tectonic map of the Rif and Betic thrust systems and the Alboran Sea. Relocated seismicity (1990–94) is also shown. Inset, map showing seismic stations used in relocating these events. A simple one-dimensional velocity model consisting of five crustal layers (with velocities (km s^{-1}) of 5.0, 6.0, 6.2, 6.7 and 7.2 and depths (km) of 7.5, 12, 18, 30 and 32) and a Moho depth of 32 km with a sub-Moho velocity of 8.2 km s^{-1} is used in relocation. Although the region is structurally very complicated, for earthquake location purposes this simple model suffices. Precision in locations is shown with error bars for each earthquake.



Alboran basin, Miocene in age^{12,13}, developed within an approximately north–south-directed compressional tectonic regime, which is a consequence of the convergence of the Africa and Iberia plates since the late Cretaceous¹⁴. Proposed models for the geodynamic evolution of the extensional Alboran basin within this compressional environment, and contemporaneous outwardly directed thrusting along the margins of the basin, have been controversial. Among the proposed models are mantle upwelling¹⁵, an Alboran block (or a micro-continent) moving towards the west and sandwiched by Iberia and Africa^{16,17}, lithospheric delamination^{18,19}, and retreating subduction²⁰. All of these models have been used to explain the kinematics of this region and the intriguing circular shape of thrusting (Fig. 1).

The Alboran Sea basin is covered by thick Miocene and younger sediments, locally reaching 7 km in thickness, resulting from early Miocene extension in the region²¹. These sediments overlay metamorphic basement that is about 16–22 Myr old^{22,23}. Extension ceased in the basin by the late Miocene when the whole area was uplifted above sea level in Messinian time¹¹. Late Miocene to recent tectonic history involves inversion²⁴, or inversion, strike-slip and extension^{11,21}. Thrusting in the Betic and Rif regions developed contemporaneously with the extension in the Alboran basin²⁵. A westward motion of the thrust front of the Gibraltar arc is also proposed²⁶. Based on regional subsidence analysis, Docherty and Banda³ proposed an eastward migration of the basin due to a possible southeasterly propagating lithospheric delamination.

Neogene volcanic activity is also observed in the Alboran Sea basin as well as in the Betic and Rif belts. Middle Miocene, and early Pliocene calc-alkaline volcanism is observed in the eastern Rif^{27,28}, eastern Betics²⁹ and in the Alboran Sea basement²⁶. Potassic volcanism is also observed in the Betics and the Alboran Sea. In addition, mantle peridotites located at the margins of the Betic and Rif systems are observed at the surface. These peridotites have been interpreted as intrusive bodies, as they show high-temperature aureoles along their contacts with the host rocks³⁰.

Seismic refraction data indicate that the Alboran Sea is characterized by a thin, continental type crust about 15 km thick, and P_n velocities in the uppermost mantle are low, $7.5\text{--}7.9 \text{ km s}^{-1}$ (refs 31, 32). The crustal thickness increases towards the coastal areas reaching about 25 km in southern Spain and 30–38 km beneath the pre-Betic areas³³. The crust also thickens towards the Rif area, reaching 25 km in the coastal areas and about 35 km beneath the pre-Rif³¹.

Seismicity and earthquake relocation

The Azores–Gibraltar seismic zone is a plate boundary separating the African and Eurasian plates. Well defined seismic activity is observed between the Azores and Gulf of Cadiz. However, seismicity between the Gulf of Cadiz and the Gibraltar strait is more diffuse³⁴. We studied the seismicity between longitudes 0° and 12° W , and latitudes 34° to 38° N . About 40 seismic stations are operating in the vicinity of the study region (Fig. 1). We have re-read all seismograms from the Moroccan seismic stations and merged our readings with the Spanish readings as reported to the USGS (in the EDR “Earthquake Data Reports”). Then we relocated all the events (about 1,200) between January 1990 and August 1994, using the hypoinverse earthquake location program³⁵ and a simple one-dimensional velocity model.

Although epicentral locations of the relocated earthquakes did not change significantly from their original locations, we observed a significant change in depth determinations. Figure 1 shows relocated earthquakes in the region, excluding all events that had total r.m.s. errors of more than 1 s. We also excluded those events that had poor depth control, by applying criteria based on maximum depth errors. For shallow events we defined a threshold of $\pm 10 \text{ km}$ in depth errors. For the intermediate-depth events, we defined a threshold based on the depth of the events. We excluded all events that had depth errors more than 20% of the calculated depth. These criteria helped us obtain a seismicity database that is reasonably accurate and does not include significant mislocations (Figs 1 and 2).

These relocated hypocentres are projected into planes oriented in a north–south direction. Depth distribution of the projected earthquakes is shown in Fig. 2 along with the calculated standard errors. A few significant patterns are recognizable from these relocated events. Both the Rif and Betic crusts are seismically active. Seismic activity is limited to the upper crust beneath the Rif region; however, both the upper and lower crust appear to be seismically active beneath the Betic region. Limited shallow seismic activity is observed beneath the Alboran Sea. Probably the most significant discovery from this relocation study is the apparent seismic gap beneath the Alboran Sea and the Gulf of Cadiz. A well defined aseismic zone between depths of about 20 km and 60 km is identified as having no, or a very limited number of, earthquakes (Fig. 2). Starting at about 60 km depth, significant intermediate-depth seismic activity is observed. This intermediate-depth seismic activity extends below 150 km depth. Intermediate-depth seismicity is also observed beneath both

continental areas of the Rif and Betic regions with maximum depths of earthquakes reaching about 200 km.

Attenuation of seismic shear waves

We use digital seismic waveform data from the Moroccan national seismic network in a qualitative mapping of regions of high- and low-attenuation zones in the Alboran Sea and the Rif and Betic regions. Seismograms (vertical component) from earthquakes located in Spain, northern Morocco and Alboran Sea show very clear evidence of strong P- and S-wave attenuation both in the crust and in the uppermost mantle of these regions. The shear waves, in particular, are strongly attenuated and in many cases are not observed within the frequency band between 2 and 10 Hz.

Examples of digital seismograms are shown in Fig. 3. Figure 3a shows seismograms of a shallow earthquake located within the Rif tectonic boundary. Stations BIT and MIF are both at approximately the same distance from the earthquake and their azimuth vary by about 180°. Seismic wave energy travelling into the Rif region (station BIT) essentially lost its entire shear wave energy along its 126 km path (Fig. 3). No simply identifiable shear arrivals are observed at or near the expected S_g time (see Fig. 3 legend for seismic wave nomenclature). However, seismic wave energy travelling to the southern station MIF clearly shows a high-frequency S_g arrival. As the travel paths to the stations are azimuthally 180° apart, the effect of source parameters on the radiation patterns and amplitudes of recorded seismic phases is minimal.

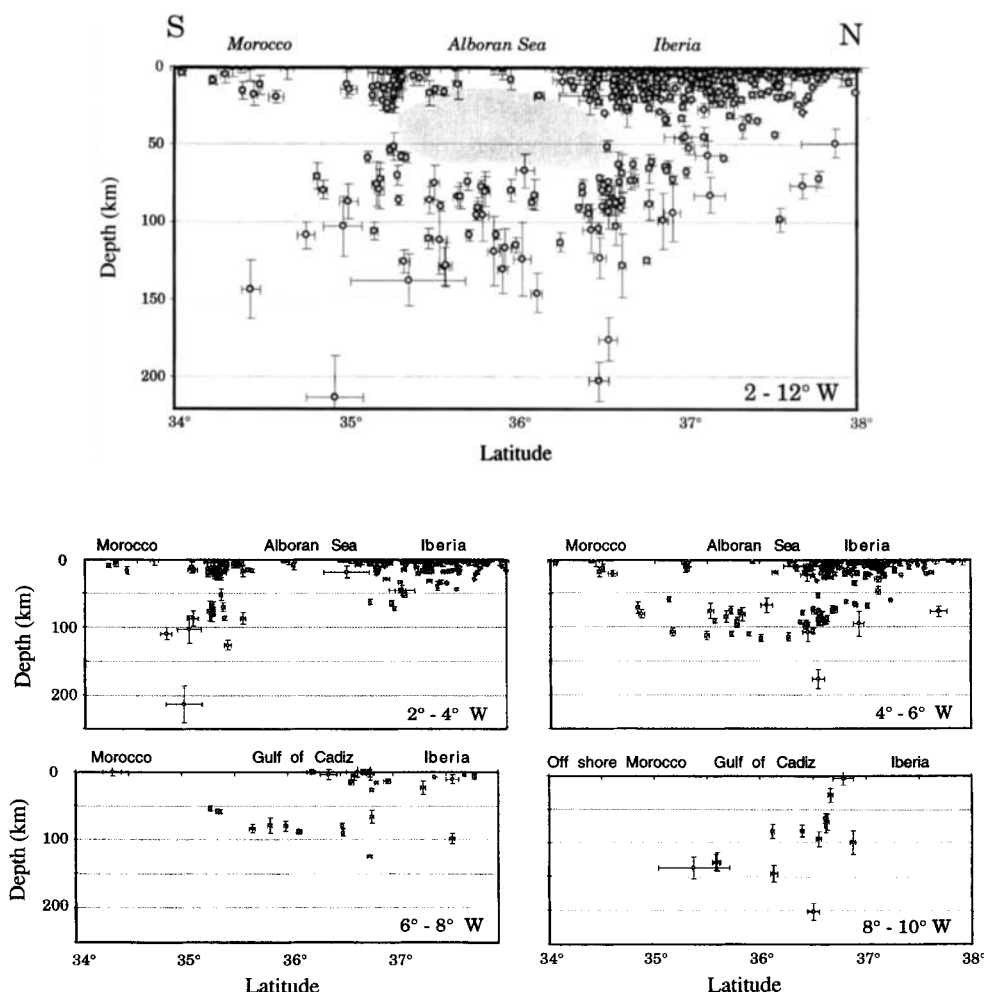
Seismograms from intermediate-depth earthquakes also show significant variations in waveform propagation. Figure 3b shows

four seismograms from an intermediate-depth earthquake located in the western Rif region. The seismograms from the Rif stations CPS and NKH not only show a lower-frequency character, but also lack significant shear wave energy. Although this may partly be due to the recordings being on vertical geophones and the rays' steeper incidence angles, a thorough analysis of several intermediate-depth earthquakes showed consistent results. Intermediate-depth earthquakes are poorly recorded by the Rif stations, sometimes with no obvious recordings at all, while stations farther to the south in the Middle Atlas and High Atlas regions do record the same events with large amplitudes and high frequencies.

These propagation anomalies are characteristic for all the earthquakes in this region. We have analysed over 50 earthquakes located in and around northern Morocco, southern Spain and the Alboran Sea regions. This qualitative analysis shows that the uppermost mantle and the lower-crustal material beneath the Rif region and the Alboran Sea are very attenuating and absorb shear wave energy even at short distances. Higher-frequency seismic waves are recorded only from local events located within a few kilometres of the stations. We cannot accurately constrain the northern boundary of these high-attenuation zones of the lower-crust and uppermost mantle as we do not have access to waveform data from southern Spain. However, the southern boundary is reasonably well mapped (Fig. 3).

The Alboran Sea, as a whole, is a region of high crustal and uppermost mantle attenuation. This attenuation zone coincides with reported Neogene volcanism in the Alboran Sea (Figs 1 and 3). The pre-Rif zone is not characterized as a high-attenuation region. The high-attenuation zone starts at approximately the Rif

FIG. 2 A general cross-section oriented north-south showing seismicity (shown in Fig. 1) projected onto a vertical plane (top), and four north-south-oriented cross-sections showing depth distribution of seismicity and lateral variations from east to west at every two-degree longitude intervals (bottom). Earthquake hypocentres with their standard errors are shown. Shaded area beneath the Alboran Sea marks the approximate location where a seismic gap is observed. Intermediate-depth seismicity is observed beneath both the Rif and the Betic belts.



thrust boundary and continues farther to the north (Figs 1 and 3). The boundary also approximately coincides with the locations of the southernmost intermediate-depth earthquakes in the Rif region (Fig. 1). The high-attenuation zone extends towards the west beneath the Atlantic Ocean where the Rif and pre-Rif thrusts continue offshore.

Three-dimensional gravity modelling

An arcuate-shaped negative Bouguer gravity anomaly approximately surrounds the Betic and Rif mountain belts (Fig. 4). Bouguer gravity anomalies as low as -155 mGal are observed in the Rif region, and about -130 mGal in the Betics. This low gravity anomaly extends into the Atlantic Ocean for a distance of about 200 km. Bonini *et al.*³⁶ interpreted these negative gravity anomalies as a manifestation of deep crustal tectonics, based on the fact that metamorphic basement in the eastern Betic is exposed and forms an antiform without any perturbation on the gravity values. Other studies, however, have interpreted the negative gravity values as a result of sediments thickened by thrust duplexes³⁷.

We performed three-dimensional gravity modelling in the southern Rif region where we have numerous drill-hole data and good seismic reflection data coverage to constrain the shallow subsurface structure and its contribution to the low Bouguer anomalies. We selected an area of about 250×100 km in the southern Rif where minimum gravity values are observed (Fig. 4A). Available drill-hole data in the western Rif provided accurate measures to metamorphic basement depth (Fig. 4B, a). The seismic reflection data were used to map the remaining area. Reflection data were tied to well data or to surface geology, and a map of depth to basement was obtained (Fig. 4B, a).

The basement geometry was then used in a three-dimensional gravity modelling³⁸ to obtain modelled Bouguer gravity values. Our model included a single layer, representing the sedimentary

rocks, over a half-space. A constant density is assumed for all sedimentary rocks. The density contrast is the only parameter that we do not have any constraints on. Initially, we assumed that the density difference between the sedimentary rocks and the metamorphic basement is -0.25 g cm⁻³. Assuming that the basement density is 2.70 g cm⁻³, this would result in an average density of 2.45 g cm⁻³ for the entire sedimentary column, which is in general composed of limestones, marls, sandstones and clastics. Figure 4 shows the results of this modelling and compares them with the observed values (Fig. 4B, b and c). The minimum value from this modelling is about -70 mGal, about half of the observed minimum value. Although the general shape appears to be quite reasonable, the low magnitude of the calculated anomaly suggests that sediments in this region cannot alone account for all the observed gravity anomalies. To search for a density variation that would yield a better fit to the magnitude of the observed anomaly without introducing density variations in the deeper crust and uppermost mantle, we performed a series of calculations by varying the density between -0.25 and -0.50 g cm⁻³ in increments of 0.05 g cm⁻³. The best fit in amplitude was obtained when the density variation was -0.50 g cm⁻³. This would correspond to an average density value of 2.20 g cm⁻³ for the entire sedimentary rock column, which is an unrealistically low average density value for an 8-km-thick sedimentary section.

Results of this three-dimensional gravity modelling suggest that sediments thickened by imbricate thrusting and foreland basins can only account for one-half of the observed gravity anomalies in the Rif region. Then, we can assume that the rest of the density anomaly is within the deep crust or even in the uppermost mantle. We do not, however, have strong constraints for the deeper structures, as we had for the sedimentary section. But our preferred interpretation is that asthenospheric material exists at subcrustal depths, as we also see evidence of high shear wave attenuation, and that this lower-density material contributes

FIG. 3 Examples of seismograms showing waveform propagation anomalies in the Rif and surrounding regions. Seismograms are band-pass filtered between 2 and 10 Hz. Station names and epicentral distances in kilometres are shown on the left of the seismograms. Stripes on the maps indicate the zone of extremely high attenuation. This zone is not well mapped in the north. P and S on the seismograms denote compressional and shear wave arrivals, respectively. Upcoming seismic waves are named as P or S. Crustal seismic phases are subscripted with "g", and refracted seismic rays along the crust-mantle (Moho) boundary are subscripted with "n". a, Seismograms from a shallow earthquake in the Rif region. Stations BIT and MIF are at the same distance at opposite directions. Strong shear wave energy in the MIF recording contrasts with lack of shear waves in the BIT recording. b, Seismograms of an intermediate-depth earthquake in the Rif region. Lack of shear waves (even at very short distances) is clear on the Rif stations' records (especially NKH). In contrast, stations CIA and OUK show strong and high frequency shear waves.

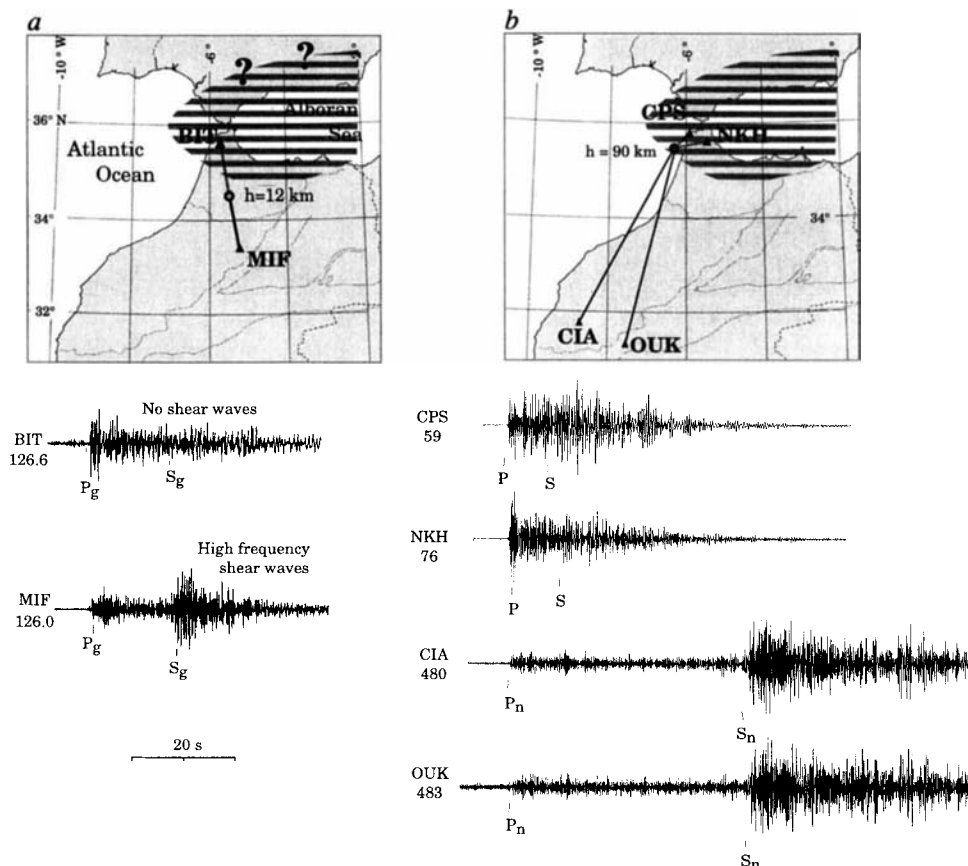
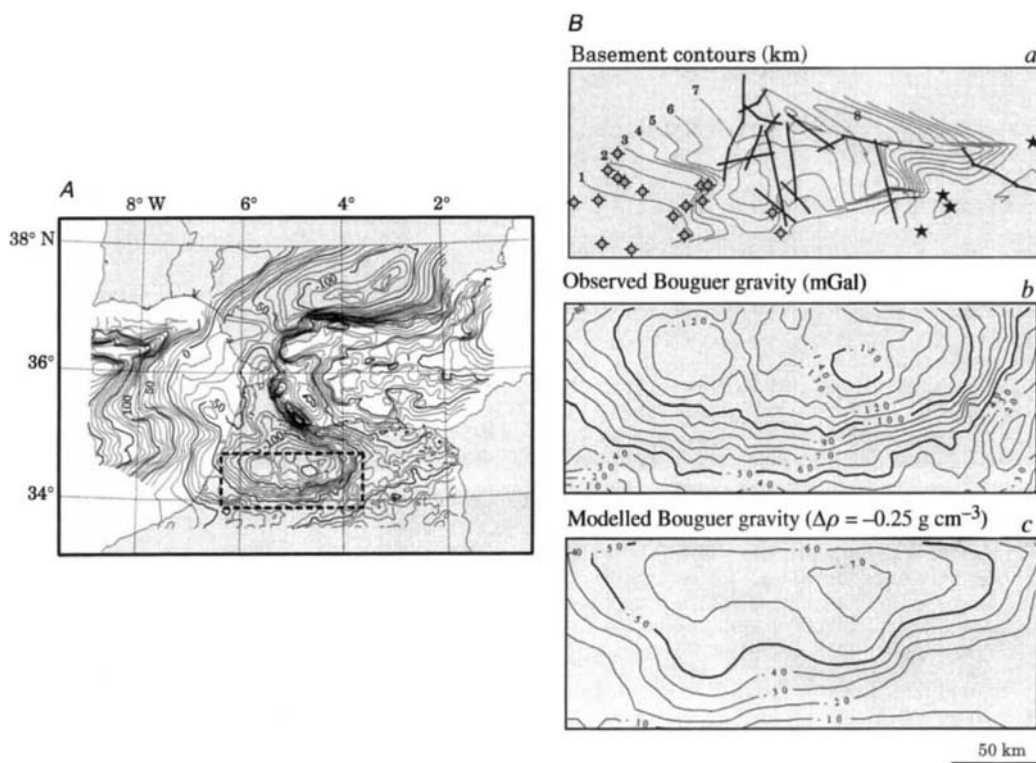


FIG 4. A, Bouguer gravity map of the Alboran Sea area and surrounding mountain belts. Box represents the area of gravity modelling. B, a, basement depth in the southern part of the Rif area. Basement contours are in kilometres. Circles with crosses represent drill-hole locations where basement was penetrated. Stars represent locations of exposed metamorphic basement. Thick black lines are locations of seismic reflection profiles used in basement mapping. B, b, Observed Bouguer gravity anomaly. B, c, calculated (three-dimensional) Bouguer gravity anomaly for the basement geometry in B, a. Clearly, sediments and the underlying basement morphology cannot alone explain the observed gravity.



to this lower Bouguer gravity anomaly surrounding the Gibraltar arc.

Manifestation of lithospheric delamination

New geophysical data and analyses presented above provide cogent evidence for continental lithospheric delamination beneath the Alboran Sea and Rif-Betic mountain belts. In addition to previous geological evidence, seismicity, low uppermost mantle seismic wave velocities and high attenuation, low gravity anomalies and high heat flow values are among the expected consequences of delamination and have been documented to exist in the Alboran Sea and surrounding regions in this study. A pronounced seismic gap with low velocities and high attenuation in the uppermost mantle that is underlain by seismically active, high-velocity, high- Q (that is, efficient in seismic wave propagation) material in the upper mantle are among the strongest geophysical evidences for lithospheric delamination. Young volcanism at the surface and high heat flow values^{39–42} in the region suggest that this zone of low velocities and high attenuation in the uppermost mantle represent asthenospheric material⁴³. We estimate from the spatial distribution of earthquake hypocentres and attenuation studies that the thickness of the asthenospheric material in the uppermost mantle is about 45–50 km beneath the Alboran Sea, and less beneath the Rif and Betic regions. The high-velocity upper mantle^{7,8} underlying this asthenospheric material is inferred to be lithospheric material delaminated from the overlying crust. Whether or not the lower crust is involved in delamination is rather controversial⁶. However, it is expected that the lower crust will be affected by the underlying asthenospheric material after delamination. Partial melting in the lower crust and/or mobilization are expected, owing to hot underlying asthenosphere. This could explain the observed high crustal shear-wave attenuation in the Rif region of Morocco (Fig. 3a).

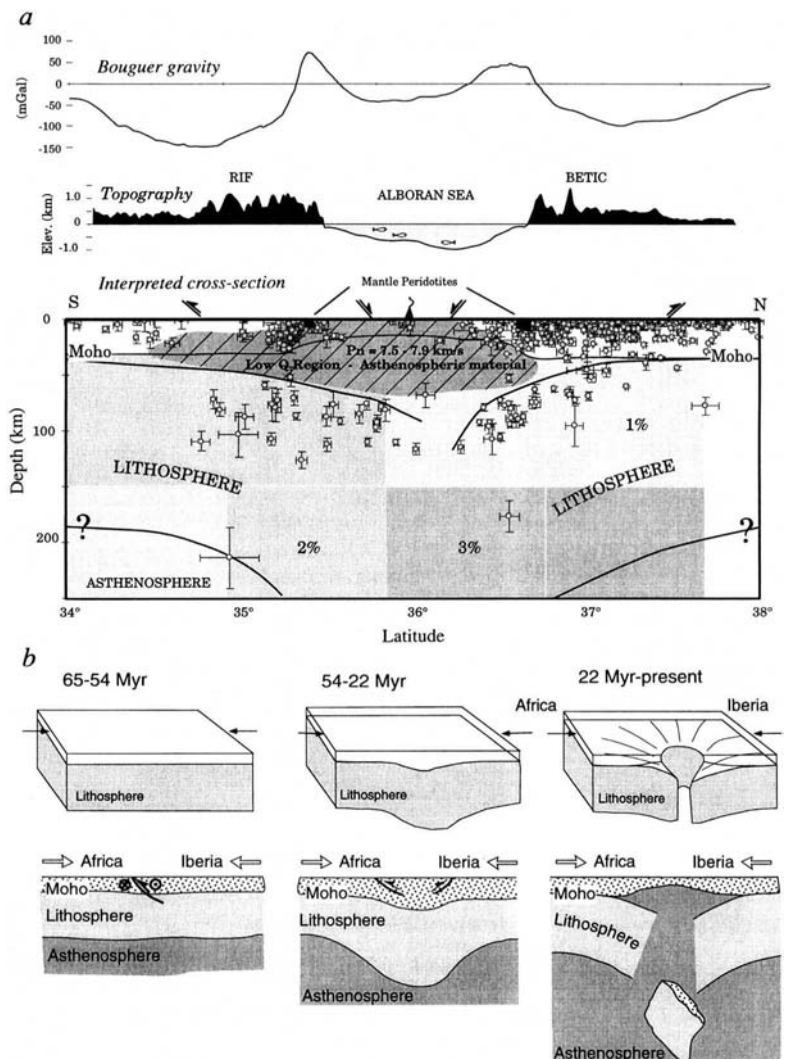
The presence of shallow asthenospheric material can also manifest itself in high heat flow values measured at the surface. Bird³ estimated that the heat from asthenospheric material at subcrustal levels takes about a few million years to reach the surface. The heat flow anomalies of this type are relatively small in

magnitude, 15–25 mW m⁻² above the background values of the region³. The reported heat flow values in the region are higher than average and fall within these limits^{39–42}.

Geophysical observations presented here, along with available geological information, eliminate several proposed geodynamic models for this region. Simple subduction is not a fitted model, because we observe subcrustal earthquakes both in the Betic and Rif regions significantly far away from any possible plate boundaries. Neither is the mantle upwelling model consistent with the high-velocity material we observe in the upper mantle. The Alboran block model is fitted to explain the circular thrusting around the Gibraltar arc, but it fails to provide any explanation of the observed structure in the uppermost mantle and especially the observed extension in the Alboran Sea basin. The model of removal of lithospheric material, on the other hand, is consistent with all of the geophysical and geological observations.

One of the remaining problems is that whether or not we can distinguish between a delamination (s.s.) model and a convective removal model. The hypocentral distribution of subcrustal earthquakes show a pattern (both beneath the Rif and Betic regions) that can be interpreted to represent nearly the top of the delaminating continental lithosphere (Fig. 5). High-frequency seismic wave propagation suggests that the delaminating lithosphere is likely to be connected to the Moroccan lithosphere in the south. Probably the same is valid for the northern part, as intermediate-depth seismicity becomes shallower northward from the Alboran Sea and continues into the Iberian lithosphere beneath the Betic region. This geometry, and the existence of asthenospheric material, as inferred from low-velocity anomalies, high seismic wave attenuation and low densities in the uppermost mantle, is more suggestive of a delamination (s.s.) model. This type of geometry (Fig. 5) also brings an understanding to the formation of the two Cenozoic basins, the Rharrb basin in Morocco and Guadalquivir basin in Spain (Fig. 1). These basins can be interpreted as flexural basins developed as a result of the sinking lithosphere⁴⁴. As the continental lithosphere sinks, it pulls down the overlying crust and creates these depressions in northern Morocco and southern Spain. The size and shape of the Gibraltar arc and the geometry of these basins are characteristic of the response of the lithosphere

FIG. 5a, Interpreted geological cross-section between the Rif and Betics. This north-south-oriented cross-section represents the inferred structures as well as topographic and gravity profiles along 5° W longitude. Seismicity shown (circles with standard errors) is between 2° and 6° W longitude. Continental lithosphere beneath both the Rif and Betics are delaminating. Base of the lithosphere in each region is not well constrained. As the lithosphere peels off from the overlying crust, low-velocity, low-Q asthenospheric material rises up and replaces the lithosphere, especially beneath the Alboran and Rif regions. Also shown are velocity variations (in per cent) obtained by teleseismic tomography⁷. b, Schematic representation of the evolution of the lithosphere in the Alboran Sea area.



when it is subjected to a point force⁴⁴. But the observed symmetry both in surface geology and subsurface structures may be considered to support the convective removal model as a better model for the region. We argue that initiation of removal of lithospheric material may have been due to thickening of the

continental lithosphere and subsequent convective removal of parts of the lithosphere, similar to what is proposed by Platt and Vissers¹⁸. It is plausible that this tectonic event in the early Miocene generated instability within the lithosphere which now is a possible driving mechanism for active delamination (s.s.). □

Received 24 August 1995; accepted 26 January 1996.

- Dewey, J. F. *Tectonics* **7**, 1123–1139 (1988).
- Nelson, K. D. *Geology* **20**, 498–502 (1992).
- Bird, P. J. *geophys. Res.* **84**, 7561–7571 (1979).
- Houssesman, G. A., McKenzie, D. P. & Molnar, P. J. *geophys. Res.* **86**, 6115–6132 (1981).
- Docherty, C. & Banda, E. *Tectonics* **14**, 804–818 (1995).
- Kay, R. & Kay, S. M. *Tectonophysics* **219**, 177–189 (1993).
- Seber, D. et al. *J. geophys. Res.* (in the press).
- Blanco, M. J. & Spakman, W. *Tectonophysics* **221**, 13–34 (1993).
- De Galdeano, C. S. *Tectonophysics* **172**, 107–119 (1990).
- Morley, C. J. *African Earth Sci.* **14**, 57–65 (1992).
- Watts, A. B., Platt, J. P. & Buhl, J. P. *Basin Res.* **5**, 153–177 (1993).
- Cloetingh, S. et al. *Geo-Mar. Lett.* **12**, 66–75 (1992).
- Campillo, A. C., Maldonado, A. & Mauffret, A. *Geo-Mar. Lett.* **12**, 165–172 (1992).
- Dewey, J., Helman, M., Turco, E., Hutton, D. & Knott, S. in *Alpine Tectonics* (eds Coward, M., Dietrich, D. & Park, R.) 265–283 (Geol. Soc. Spec. Publ. 45, Oxford, 1989).
- Weijermars, R. *Phys. Earth planet. Inter.* **37**, 135–148 (1985).
- Duran-Delga, M. & Oliver, P. in *The Atlas System of Morocco* (ed. Jacobshagen, V.) 465–480 (Springer, Berlin, 1988).
- Andrieux, J., Fontbote, & Mattauer, M. *Bull. Soc. géol. Fr.* **27**, 115–118 (1971).
- Platt, J. P. & Vissers, R. L. M. *Geology* **17**, 540–543 (1989).
- Vissers, R. L. M., Platt, J. P. & van der Wal, D. *Tectonics* **12**, 804–818 (1995).
- Royden, L. *Tectonics* **12**, 629–638 (1993).
- Comas, M., Garcia-Duenas, C. V. & Jurado, M. J. *Geo-Mar. Lett.* **12**, 157–164 (1992).
- Steiger, R. H. & Frick, U. *Init. Rep. DSDP Leg 13(2)*, 762 (1973).
- Docherty, C. & Banda, E. *Geo-Mar. Lett.* **12**, 82–87 (1992).
- Bourgeois, J., Mauffret, A., Ammar, A. & Demnati, A. *Geo-Mar. Lett.* **12**, 117–122 (1992).
- Hovath, F. & Berckhemer, H. in *Alpine Mediterranean Geodynamics* (eds Berckhemer, H. & Hsu, K.) Vol. 7, 141–173 (Geodynamics Ser., American Geophysical Union, Washington DC, 1982).

- Garcia-Duenas, V., Balanya, J. C. & Martinez-Martinez, J. M. *Geo-Mar. Lett.* **12**, 88–95 (1992).
- Ait Brahim, L. thesis, Univ. Mohammed V, Rabat, Morocco (1991).
- Ait Brahim, L. & Chotin, P. J. *African Earth Sci.* **11**, 273–280 (1990).
- Arana, V. & Vegas, R. *Tectonophysics* **24**, 197–212 (1974).
- Doblas, M. & Oyarzun, R. *Earth planet. Sci. Lett.* **93**, 76–84 (1989).
- Hatzfeld, D. & Ben Sari, D. *Bull. Soc. géol. Fr.* **7**, 749–756 (1977).
- Working Group for the Deep Seismic Sounding in the Alboran Sea *Pure appl. Geophys.* **116**, 167–180 (1978).
- Banda, E. *Phys. Earth planet. Inter.* **51**, 222–225 (1988).
- Burom, E., Udias, A. & Colombas, M. A. *Tectonophysics* **152**, 89–118 (1988).
- Klein, F. W. *Hypercenter Location Program, HYPOINVERSE* (US open file 178-694, US Geological Survey, Reston, Virginia, 1978).
- Bonini, W. E., Loomis, T. P. & Robertson, J. D. J. *geophys. Res.* **78**, 1372–1382 (1973).
- Demnati, A. *Bulletino di Geofisica Teorica ed Applicata* **14**, 203–236 (1972).
- Gotze, H. J. & Lahmeyer, B. *Geophysics* **53**, 1096–1108 (1988).
- Polyak, B. G. et al. *Ann. Geofisica* **13**, C18 (1994).
- Rimi, A. *Geothermics* **19**, 443–454 (1990).
- Cermak, V. in *Terrestrial Heat Flow in Europe* (eds Cermak, L. & Rybach, L.) (Springer, Berlin, 1979).
- Albert-Beltran, J. F. in *Terrestrial Heat Flow in Europe* (eds Cermak, L. & Rybach, L.) 261–266 (Springer, Berlin, 1979).
- Molnar, P. & Oliver, J. J. *geophys. Res.* **74**, 2648–2682 (1969).
- Turcotte, D. & Schubert, G. *Geodynamics, Applications of Continuum Physics to Geological Problems* 450 (Wiley, New York, 1982).

ACKNOWLEDGEMENTS. We thank C. Docherty, R. K. Litak, D. Nelson, J. Oliver and E. Sandvol for their reviews of earlier versions of this paper, and D. Turcotte for discussions. This work was supported by the US NSF.

Control of inflorescence architecture in *Antirrhinum*

Desmond Bradley, Rosemary Carpenter, Lucy Copsey, Coral Vincent, Steven Rothstein* & Enrico Coen

Genetics Department, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK

Flowering plants exhibit two types of inflorescence architecture: determinate and indeterminate. The *centroradialis* mutation causes the normally indeterminate inflorescence of *Antirrhinum* to terminate in a flower. We show that *centroradialis* is expressed in the inflorescence apex a few days after floral induction, and interacts with the floral-meristem-identity gene *floricaula* to regulate flower position and morphology. The protein CEN is similar to animal proteins that associate with lipids and GTP-binding proteins. We propose a model for how different inflorescence structures may arise through the action and evolution of *centroradialis*.

PLANTS exhibit highly varied architectures, from simple stems terminating in a flower to large, highly branched structures¹. This variation is largely dependent upon where and when flowers are made²⁻⁴. In plants with determinate inflorescences the main axis of growth terminates in a flower, and further flowers are produced by branching events that occur below this. In species with indeterminate inflorescences, the main axis is prevented from forming a flower and can therefore grow indefinitely, allowing more flowers to form on an elongated stem (illustrated by the wild-type *Antirrhinum* inflorescence in Fig. 1). Evolution is thought to have proceeded from determinate to indeterminate growth, presumably by recruiting genes that repress the formation of terminal flowers^{5,6}. The *centroradialis* (*cen*) mutation of *Antirrhinum* causes the inflorescence to terminate in a flower, conferring the determinate condition (Fig. 1)^{7,8}. To understand how the formation and repression of flowers are coordinated to give a particular plant architecture, we have analysed the mechanism by which *cen* acts.

In *Antirrhinum*, the apical meristem of the primary shoot first undergoes a vegetative phase, generating leaf primordia on its flanks that bear secondary shoot meristems in their axils (the region between the stem and leaf). Under the appropriate developmental and environmental signals, such as long day length, the apical meristem is converted to an inflorescence meristem^{2,9}, which produces modified leaf primordia (bracts) with floral meristems in their axils. The apical meristem maintains its inflorescence identity throughout the remainder of its lifetime, generating flowers in axillary positions (Fig. 1). This pattern is reflected in the expression of floral-meristem-identity genes such as *floricaula* (*flo*) and *squamosa* (*squa*), which are expressed in floral meristems and their subtending bracts but not in the apical meristem^{9-11,33}. Floral meristems give rise to four concentric whorls of organs: five sepals outermost, surrounding five petals which are united for part of their length, followed by four stamens and the central two fused carpels. *Antirrhinum* flowers are zygomorphic, having a single plane of symmetry^{12,13}.

The *cen* mutant of *Antirrhinum* differs from wild type in several key respects. (1) The mutant produces a terminal flower, converting the inflorescence from indeterminate to determinate (Fig. 1). Consequently, the architecture is changed to a shorter, more bushy plant, as shoots cannot grow indefinitely. (2) About 10 axillary flowers are made below the terminal flower, indicating that there may be a delay in switching from an inflorescence to a

floral meristem (see legend to Fig. 1). (3) The terminal floral meristem is developmentally more advanced than the axillary flowers below it. (4) Unlike axillary flowers, organ numbers and their arrangement (phyllotaxy) are very variable in terminal flowers (Fig. 1). (5) The terminal flower is usually radially symmetrical, with all petals resembling the ventral (lowest) petal of axillary flowers. Together, these five aspects of the *cen* mutant phenotype indicate a primary role for *cen* in the programming of the apical meristem and key interactions with floral-meristem-identity genes. Mutants have also been described in *Digitalis* and *Arabidopsis* that convert an indeterminate to a determinate meristem with the production of a terminal flower¹⁴⁻¹⁶. The



FIG. 1 Inflorescences of wild-type *Antirrhinum* (left) and the *cen* mutant (right). The number of axillary flowers below the terminal flower ranged from 6 to 16 for the main stem, and 4 to 15 for secondary branches; 439 inflorescences were scored from 11 different families carrying the *cen*-594, *cen*-663 or *cen*-665 alleles. Although terminal flowers often had the same number of organs as axillary flowers, there was variation, from 3 to 8 sepals and 4 to 9 petals in the terminal flowers of secondary branches. The number of stamens also varied and these were often petaloid. Greater variation was observed in the terminal flowers of the main stem than in those of secondary branches.

* Present address: Department of Molecular Biology and Genetics, University of Guelph, Guelph, Ontario, N1G 2W1 Canada.

other aspects of the *cen* mutant phenotype are also seen in the *cen*-like mutants of these species. An additional feature of the *Arabidopsis* mutant, *terminal flower (tff)*, is that it flowers earlier than wild-type plants.

Here we describe the molecular and morphological analysis of *cen* mutants, showing when and where *cen* acts in relation to flowering. The protein CEN has some similarities to a family of phosphatidylethanolamine-binding proteins (PBP) found in animals; PBPs are thought to interact with membranes and possibly GTP-binding proteins^{17,18}. In wild type, *cen* is most strongly induced in the region just below the shoot apical meristem, and may act cell-non-autonomously to prevent *flo* expression in the apex. By analysis of *flo* and *cen* induction, we suggest a model that explains the architectures and floral morphologies of both the wild type and *cen* mutant of *Antirrhinum*.

Morphology and development of *cen* mutants

To understand how and when *cen* acts in relation to inflorescence and floral induction, we analysed the morphological development of *cen* mutant inflorescences by using scanning electron microscopy (SEM). Three aspects of the terminal flower of *cen* mutants were investigated: phyllotaxy, organ number, and developmental age in comparison to axillary flowers. SEM analysis of wild-type plants has defined several morphologically distinct stages of floral meristem development¹¹. Stage 0 is marked by the presence of small bract primordia arising on the flank of the apex. At stage 1, floral meristems comprise an eye-shaped group of cells lying in the axil of a bract. At stage 2, floral meristems are raised up to give a loaf-like appearance. By late stage 2, the floral meristem is

distinguishable from a vegetative meristem by its larger size and rounder shape. By stage 3, the pentagonal shape of the floral meristem is clear, and by stage 4 the sepal primordia are visible. Plants mutant for *cen* were grown under short days to maintain vegetative growth, and were switched to long days to induce inflorescence and floral development. Apices of plants collected on 0, 4, 6, 7, 8 or 10 days after induction were analysed by SEM (Fig. 2). One set of plants was left to mature in long days and produced 10 (± 1) axillary flowers below the terminal flower.

Under SEM, *cen* mutant plants collected 4 days after induction looked similar to plants at day 0, with the youngest bract primordia arising sequentially in a spiral on the flanks of the inflorescence meristem (Fig. 2). By day 6, axillary meristems could be distinguished as either vegetative or floral (Fig. 2). The lowest floral meristem was labelled flower 1, and all meristems above it were numbered sequentially. Flower 1 was at about stage 2–3 on day 6 and had about 11–12 primordia above it. Because only 10 flowers were eventually produced, the youngest one or two primordia at the apex are thought to have become the sepals of the terminal flower. By day 7, a morphological change was clearly detectable in the terminal meristem that distinguished it from a normal inflorescence apex. At this stage, all ten axillary floral meristems could be identified, each in the axil of a bract. The terminal meristem was more advanced than any of the axillary floral meristems, with larger sepal and petal primordia visible; the terminal meristem was at about stage 4, whereas flower 1 was at stage 3. The sepals of the terminal meristem did not usually arise simultaneously in a whorl, as seen in axillary meristems (for example, flower 1), but were most often in a spiral, sometimes as two whorls of three, or

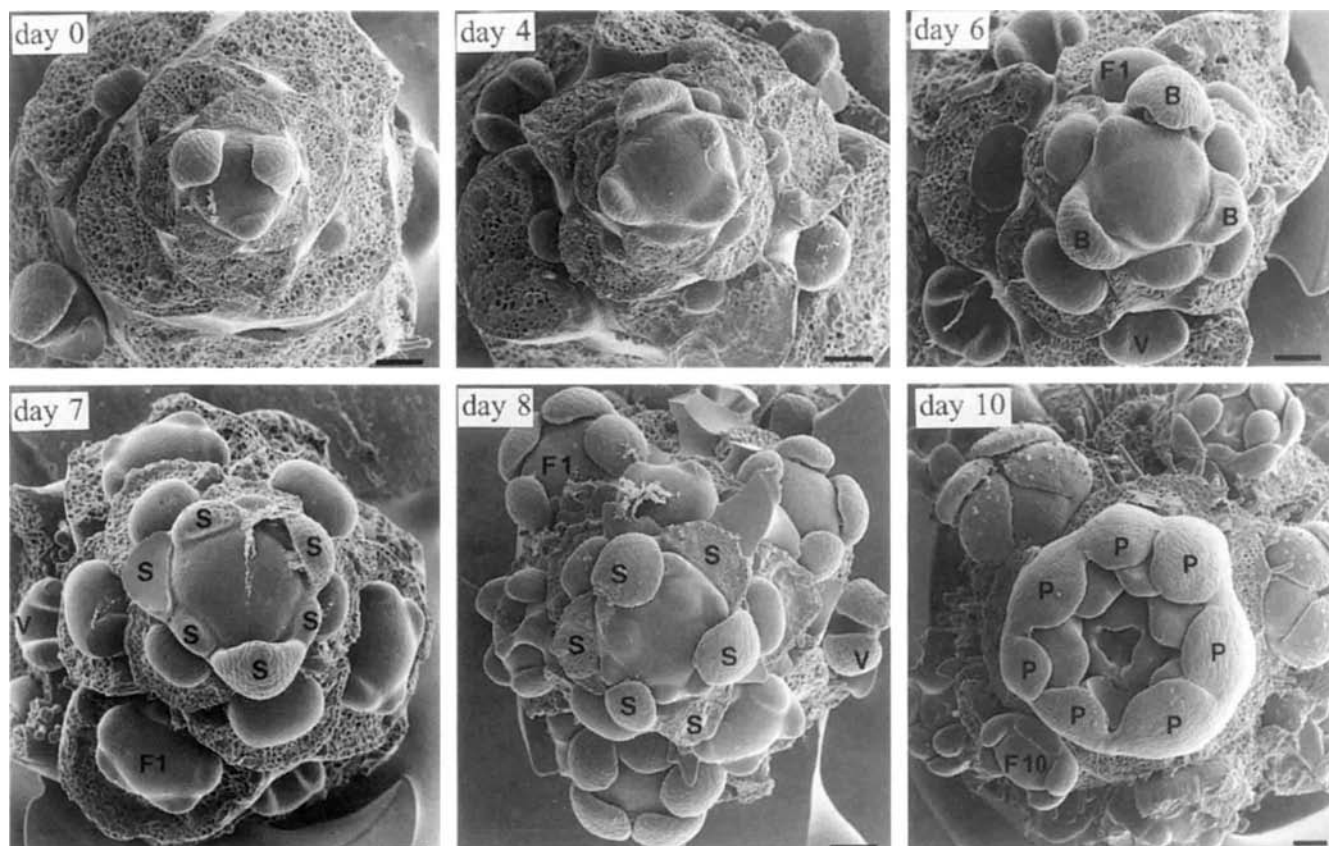


FIG. 2 Inflorescence development in the *cen* mutant of *Antirrhinum*. SEMs were made of inflorescence apices, collected on day 0, 4, 6, 7, 8 or 10 after plants were switched to long days to induce flowering. Flowers were numbered sequentially, flower 1 (F1) being the lowest and F10 residing just below the terminal flower. The sepals (S) and petals (P) are indicated for the terminal flower, some of which removed to reveal changes occurring

in the inner whorls. Bracts (B) were also removed from most axillary flowers to reveal the developing meristem. Scale bars, 100 μ m.

METHODS. Induction of flowering used plants carrying the *cen*-663 allele (J1.663) and was performed as described in Fig. 3. SEMs were made from plastic replicas²⁸.

even as a whorl of four together with one or two individual primordia. By day 8, the terminal floral meristem had clear sepal, petal and stamen primordia, whereas the petal and stamen primordia were only just apparent in the oldest axillary floral meristem. The terminal meristem usually had six sepals, most commonly arranged in a spiral.

The terminal flowers of plants analysed under SEM showed some variation in organ numbers, reflecting differences seen in plants left to mature; 6 ± 1 sepals (range 4–7). Although petal numbers also showed a small variation, they usually arose in a whorl (Fig. 2). At later stages of development, the terminal floral meristem often had five or six stamens internal to the petals, and usually two fused carpels, although three could sometimes occur.

The terminal floral meristem, therefore, behaved similarly to an axillary floral meristem; it developed at about the same rate as the axillary floral meristems (always being about one developmental stage ahead of flower 1 at any given time), and it usually made a similar number of organs to the axillary flower.

Expression of *flo* in the *cen* mutant

Two explanations might account for the more advanced nature of the terminal meristem: either it was initiated slightly earlier than flower 1, or it was recruited to form a flower at a more advanced stage of development. These two possibilities were tested by analysing *flo* expression. In wild-type apices, *flo* is activated in the youngest primordia flanking the apical meristem (stage 0),

within 1–2 days of floral induction, and is an early marker for commitment to floral meristem identity³³. We therefore compared *flo* induction in the *cen* mutant with wild type to determine its time of activation in the apex, and whether its expression domain or time of induction were consistent with the *cen* phenotype.

Wild-type and *cen* mutant plants were kept in short-day conditions to maintain vegetative growth before switching to long days to induce flowering. Apices were collected on days 0, 1, 2, 3, 4, 6, 7, 8 or 10 (long days) for RNA *in situ* hybridization (Fig. 3). Similar to wild-type plants, *flo* was induced after 1–2 long days in *cen* mutants in the youngest bract primordia arising from the flanks of the apical meristem (data not shown). Between days 0 and 4, about 5–7 bract primordia were initiated that expressed *flo*, and floral meristems at about stage 1 could be distinguished in the axils of older bracts (Fig. 3a). No expression of *flo* was observed in the apical meristem of *cen* mutants until day 6, after which *flo* expression was observed throughout the apical meristem (Fig. 3b). Wild-type control plants studied at this or any later time point after induction did not show any *flo* expression in the apical meristem (Fig. 3c; an example of wild type 6 days after induction). The ectopic expression of *flo* in the *cen* mutant shows that *cen* normally prevents *flo* expression in the apex.

According to SEM of *cen* mutants, the youngest primordia flanking the apex by day 6 would form the sepals of the terminal flower, so that it was similar to about a stage 3–4 floral meristem. Consistent with this, the expression pattern of *flo* in the apical

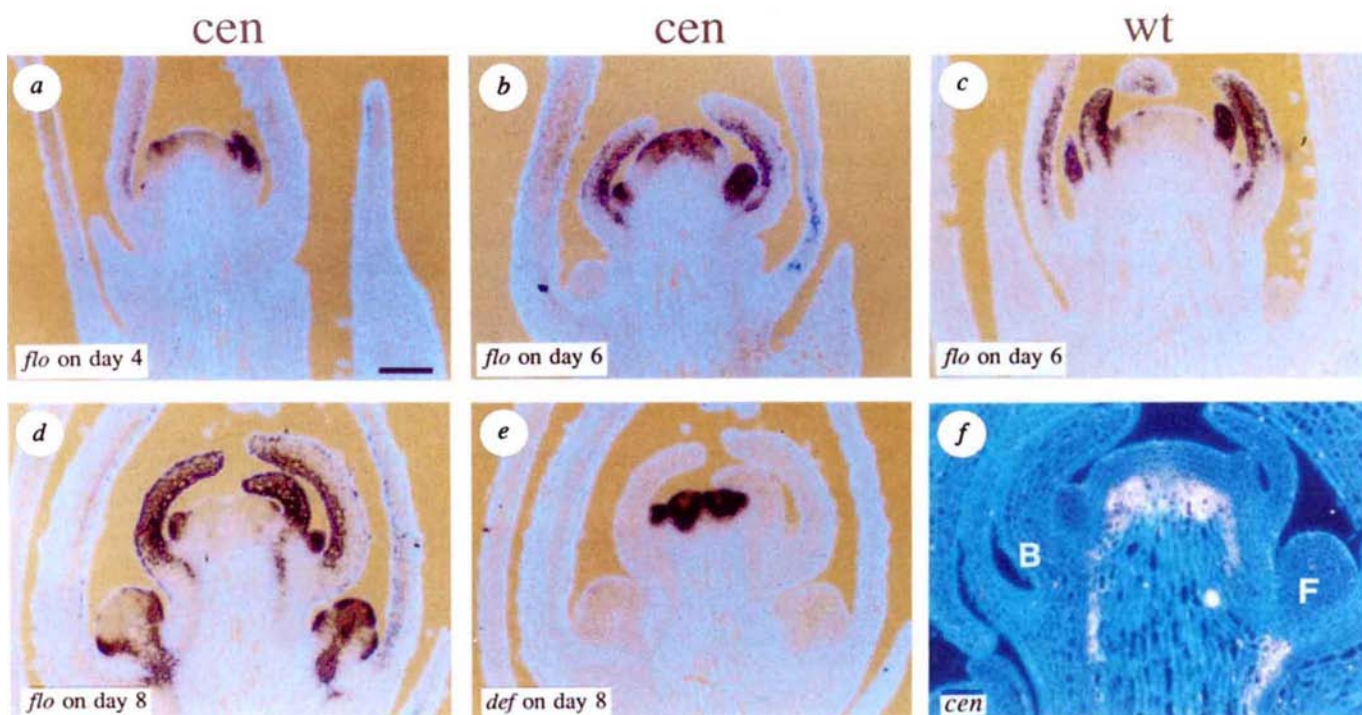


FIG. 3 RNA *in situ* hybridizations with *flo*, *def* and *cen*. a, b and d, Expression of *flo* in the *cen* mutant of *Antirrhinum* collected on day 4, 6 or 8, respectively, after floral induction. (Note that the more domed shape of the apex in b, compared to a or d, was due to the plane of section.) c, Expression of *flo* in a wild-type (wt) inflorescence on day 6. e, Expression of *def* in the *cen* mutant on day 8. Longitudinal sections of inflorescences were probed with digoxigenin-labelled antisense RNA. When viewed under light-field the RNA signal is purple on a white tissue background. Plants were grown under short-day conditions to prevent inflorescence development and then switched to long days to promote flowering. f, Expression of *cen* in wild-type (wt). Apices of mature wild-type inflorescences were collected from the field and probed for *cen* expression. When viewed under dark-field the RNA signal is a bright pale pink/orange colour on a blue tissue background. All sections are at the same scale; scale bar, 100 μ m. Note

that apices of *cen* mutants were morphologically similar to wild type until about day 6–7 (see Fig. 2).

METHODS. For a–e, *cen* mutant plants (JL.663) or wild type (JL.2) were grown under environmentally controlled conditions at about 200 microEinstein (μ E) $s^{-1} m^{-2}$ and 25 °C for about 40 short days (8 h light:16 h dark) before changing to long days (16 L:8 D). Plants were collected on the day before the change day (day 0) or on subsequent days after exposure to the extended day length. The *flo* and *def* probes and methods for digoxigenin labelling of RNA probes, tissue preparation and *in situ* hybridization were as described previously^{9,20}. An internal *AccI*–*RsaI* fragment of the partial *cen* cDNA pJAM2020 was subcloned in to Bluescript vector KS+ and used to generate antisense and sense control probes using T3 and T7 polymerases.

meristem on day 6 was typical of a stage 3–4 floral meristem. To determine the developmental stage of the terminal meristem, consecutive sections of apices were probed for expression of *flo* or the floral-organ-identity gene *deficiens* (*def*), which is first expressed at about stage 3–4 in axillary meristems^{19,20}. By day 7, *flo* expression was maintained throughout the apical meristem, while the *def* gene was expressed in the centre of the terminal meristem and was not detected in any axillary meristem (data not shown). After 8 days, the pattern of *flo* expression in the terminal meristem was typical of a stage 4–5 floral meristem (Fig. 3d). By 8 days, the expression of *def* was more extensive in the apical meristem and just visible in the oldest axillary floral meristem, flower 1 (Fig. 3e). These results suggest that the apical meristem of the *cen* mutant was recruited at a late stage (stage 2–3) to floral development, thus appearing more advanced than any axillary meristem.

Isolation of *cen*

To study the molecular action of *cen* and its relationship to inflorescence development, we isolated *cen* by transposon-tagging. Transposon-induced mutations of *cen* were obtained by crossing transposon-active lines to a line carrying the classical *cen* allele, *cen*-594 (refs 7, 8). By this strategy, 8,339 F₁ plants were screened and three new alleles of *cen* (*cen*-663, *cen*-665 and *cen*-666) were obtained. A Tam6 probe identified a 6.0-kb band that was uniquely present in *cen*-663 and linked to the *cen* phenotype. The DNA fragment was cloned, mapped and a flanking region used to probe DNA from different *cen* mutants and wild-type siblings, all heterozygous for *cen*-594 (Fig. 4). The restriction patterns were all consistent with the probe being part of the *cen* locus (Fig. 4). To confirm this, homozygous *cen*-594, *cen*-663 and *cen*-665 were grown and self-pollinated at 15 °C to induce transposon excision²¹. A small proportion of progeny from these self-pollinated plants were found with a revertant wild-type phenotype, indicating that these alleles were genetically unstable, as expected if they were caused by a transposon insertion. The revertants in each case had a restored wild-type band of 6.5 kb, presumably due to transposon excision (Fig. 4).

Overlapping clones from a wild-type genomic library were isolated using the 2-kb flanking probe and used to construct a map of the *cen* region (Fig. 5a). A wild-type 6.5-kb *Eco*RI

fragment was subcloned and fully sequenced. The insertions causing the different *cen* alleles were first mapped by genomic DNA blots. This indicated that an *Acc*I–*Eco*RI fragment at the right-hand end of the 6.5-kb *Eco*RI fragment was critical to *cen* function. However, when this and other fragments of the 6.5-kb *Eco*RI clone were used to probe a complementary DNA-library made from poly(A) RNA from inflorescences of wild-type *Antirrhinum*, no hybridizing clones were detected, presumably due to the low abundance of *cen* RNA. About 200 bp flanking the *cen*-663 allele were sequenced, and several oligonucleotides based on this sequence were used in reverse transcription–polymerase chain reaction (RT–PCR) on total RNA from young inflorescences of wild-type *Antirrhinum* or *cen* mutants. Only oligonucleotides pointing left to right, 5' to 3' in Fig. 5, gave a PCR product (in combination with an oligonucleotide designed to hybridize with the poly(A) tail). No PCR product was detected using RNA from the *cen* mutants.

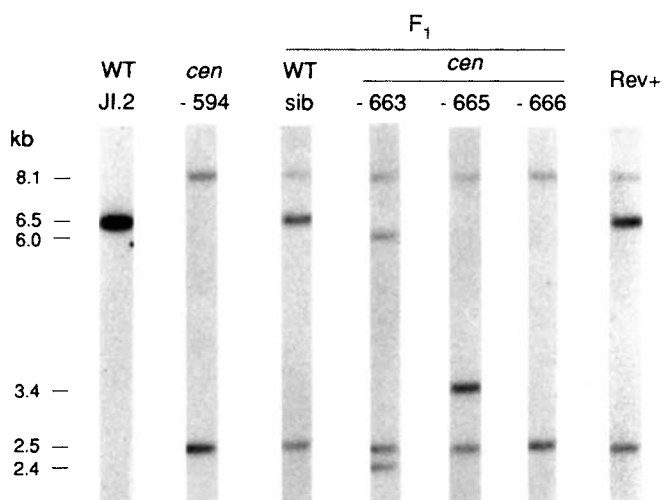
A PCR cDNA was isolated, sequenced and compared to the genomic region to determine the intron–exon boundaries. The 5' end of the *cen* mRNA was determined by 5' RT–PCR, and the complete cDNA and open reading frame (ORF) determined (Fig. 5b, c). The gene consisted of four exons comprising about 980 bp. The ORF could encode a 181-amino acid protein of relative molecular mass 20,300 (*M_r* 20.3K). No obvious structural motifs were found in the protein CEN using standard programs²². Using a conserved oligonucleotide to the CACTA end of a family of transposons in *Antirrhinum*, in combination with oligonucleotides to the *cen* region, the different *cen* alleles were mapped precisely (Fig. 5b). Two alleles, *cen*-663 and *cen*-665, were caused by Tam insertions in the first exon, and *cen*-594 was caused by a Tam4 insertion in the second intron.

Proteins similar to CEN

Database searches indicated that the CEN protein had highest similarity to a family of phosphatidylethanolamine-binding proteins present in animals, isolated based on their affinities to organic anions or lipids, or through their association with membrane–protein complexes (Fig. 5d)^{17,18}. The similarity extended throughout the proteins, and a potential nucleotide-binding region was partly conserved (CEN residues 116–132). Biochemical evidence suggests that these proteins may also

FIG. 4 Genomic DNA blot. DNA from the wild-type *Antirrhinum* progenitor line JI.2 (WT), the *cen* allele (*cen*-594) from the Gatersleben collection, Germany³⁴, and three new *cen* alleles (*cen*-663, *cen*-665, and *cen*-666) identified in the F₁ population arising from a cross between mutagenized JI.2 plants crossed to *cen*-594, were digested with *Eco*RI, blotted and probed with the blanking region of pJAM2017 (see Fig. 5 for map). A wild-type F₁ sibling (WT sib) arising from the mutagenesis and a wild-type revertant (Rev+) arising from the *cen*-594 allele were treated similarly. The wild-type progenitor gave a 6.5-kb band, whereas the *cen*-594 mutant, used as one parent in the tagging experiment, gave 8.1-kb and 2.5-kb bands. As expected, wild-type F₁ plants (*cen*-594/*Cen*⁺) had all three bands (see WT sib in Fig. 4). The three F₁ *cen* mutants had lost the wild-type 6.5-kb band, but all mutants had the 8.1-kb and 2.5-kb bands derived from the *cen*-594 parent. In *cen*-663, a 6.0-kb band was detected, equivalent to that cloned (variable levels of a 2.4-kb band were also observed; see legend to Fig. 5). The *cen*-665 mutant had a new band of 3.4 kb. The *cen*-666 allele had neither the wild-type nor any new band, and was never obtained in a homozygous state, suggesting that it carried a large and possibly lethal deletion.

METHODS. The original *cen* allele, *cen*-594, was obtained from Gatersleben^{7,8,34}. A line derived from wild-type JI.2 was grown at 15 °C and then crossed with *cen*-594. F₁ progeny (8, 339) from these crosses were grown and three new *cen* alleles, *cen*-663, *cen*-665 and *cen*-666, were obtained. These F₁ plants and three wild-type siblings from each family were maintained as cuttings for self-pollinating and crossing. Genomic DNA from each F₁ *cen* mutant and three wild-type siblings were probed with fragments of transposons Tam1–Tam8 to find any new insertion unique to each allele. A second screen of these new insertions was made on populations segregating for each *cen* allele to establish linkage with the *cen* mutant phenotype. A Tam3 insertion provided an independent restriction fragment



length polymorphism (RFLP) marker for distinguishing plants carrying *cen*-594 from the other alleles (data not shown). The individual *cen* alleles were obtained by backcrossing the F₁ *cen* mutants to wild type and selecting progeny, using the RFLP marker, that carried the new alleles. These were self-pollinated, and the segregating progeny were tested for linkage of new insertions with the *cen* phenotype. Growth of plants and methods for DNA extraction and blot analysis were as described previously^{12,20}.

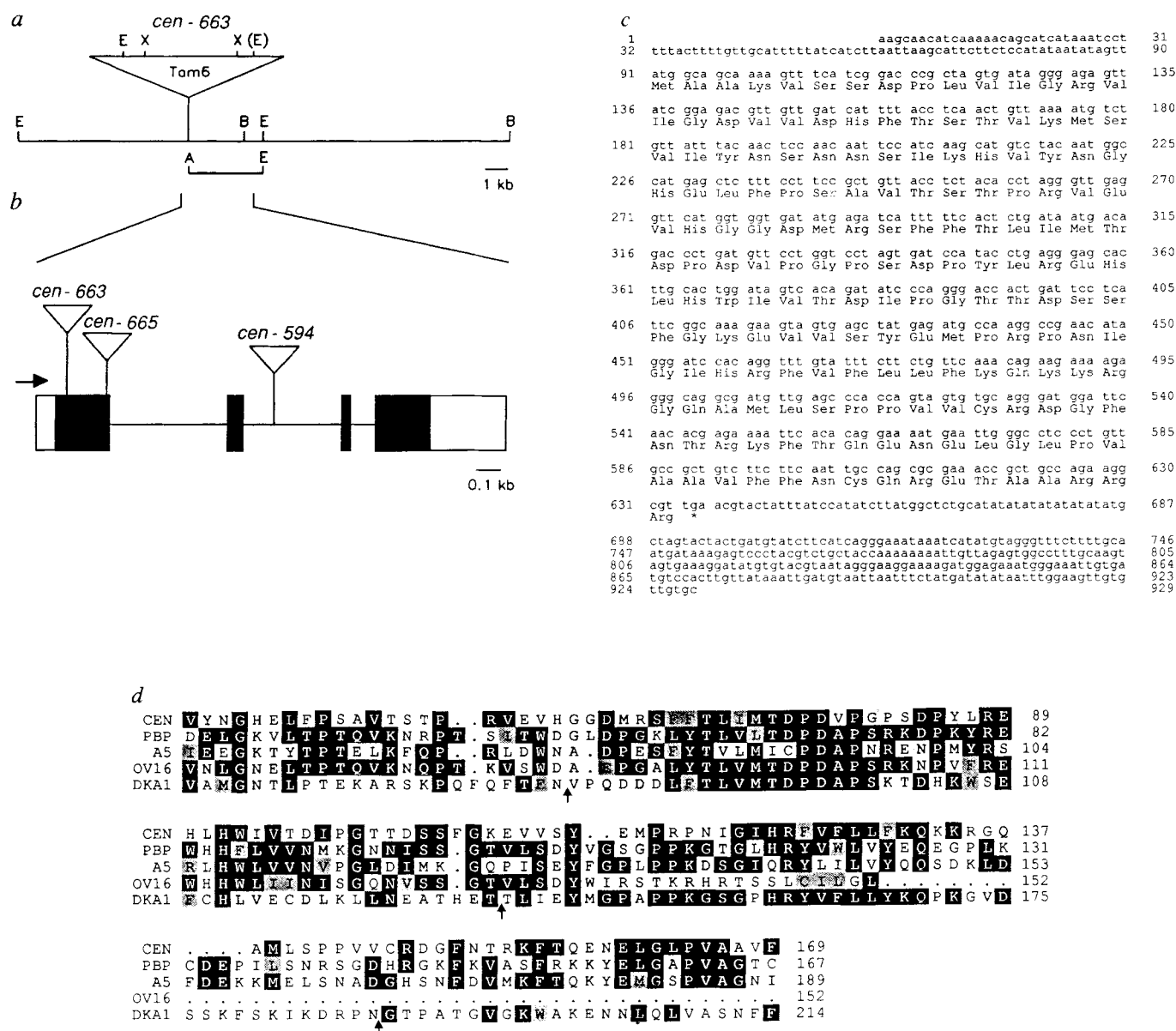


FIG. 5 The *cen* locus. **a**, Map of the *cen* genomic region carrying the *cen*-663 allele. The insertion site of the transposon Tam6 is shown, with EcoRI (E), XbaI (X) and BamHI (B) sites. An internal Tam6 XbaI fragment was used in screening to isolate the 6.0-kb EcoRI fragment (cloned as pJAM2017) that segregated with the *cen* phenotype in plants carrying *cen*-663. An EcoRI site that only partly cut in genomic digests allowed isolation of pJAM2017. **b**, Structure of the *cen* gene and the insertion sites of the transposon-generated alleles. Exons are represented by boxes, shaded for coding and open for non-coding, with introns represented by horizontal lines. Triangles on vertical lines show the transposon insertion sites. The arrow shows the direction of transcription. **c**, Nucleotide sequence of the *cen* cDNA compiled from 5' and 3' RT-PCR products and comparison with the genomic sequence. The deduced amino-acid sequence from the longest ORF is shown below. **d**, Similar sequences to CEN using BLAST²⁹. The amino-acid sequence (one-letter code) is given for regions (indicated by the amino-acid numbers on the right) of the deduced protein gene products of the four highest scores: phosphatidylethanolamine-binding protein (PBP) of rats^{17,18}, antennal 5 protein (A5) of *Drosophila*, 16K antigen from *Onchocerca volvulus* (OV16), and the *dkal* gene product of yeast (DKA1)²³⁻²⁶. The small vertical arrows indicate three regions of the DKA1 sequence which have been left out to simplify the comparison.

METHODS. The 6.0-kb EcoRI fragment identified in *cen*-663 with Tam6 was isolated, ligated to lambda gt10 arms, packaged *in vitro* (Amersham RPN1713, N334L), and a library of about 150,000 recombinants was screened with a 4-kb XbaI fragment of Tam6 (ref. 20). One positive was

isolated, subcloned as pJAM2017, which when mapped revealed an internal EcoRI site. The region flanking Tam6, a 2-kb AccI-EcoRI fragment shown as a thicker line beneath the locus in **a**, was used to screen a lambda EMBL4 library of wild-type *Antirrhinum* DNA (provided by H. Sommer). From about 500,000 recombinants, 7 overlapping clones were isolated, which were used to construct a map of the genomic region. Insertion sites were determined by sequencing products from PCR on genomic DNA of each allele, with oligonucleotides to *cen* and a conserved oligonucleotide to the CACTA family of transposable elements³⁰. The 6.5-kb EcoRI fragment, pJAM2018, contained the insertion sites of all alleles but did not identify any clones when used as a probe against a cDNA library constructed from RNA of young inflorescences of wild-type *Antirrhinum*³¹. Therefore, 200 bp flanking the *cen*-663 allele were sequenced and oligonucleotides designed in both directions for RT-PCR on RNA from wild-type and *cen* mutant inflorescences^{30,32}. This identified a cDNA, subcloned as pJAM2020, originating from the region flanking the insertion in *cen*-663 which was not expressed in each of the alleles. Both the genomic and cDNA clones were sequenced. The 5' end of the *cen* mRNA was determined using the 5' RACE system (GibcoBRL). Although a 65-bp longer 5' PCR product was found, most evidence indicated that the transcript shown is likely to be the major *cen* mRNA. First, competitive PCR suggested that the larger transcript was at least 10-fold lower in abundance; second, the proteins with similarity to CEN (including that from the cDNA isolated from *Arabidopsis*) do not have such a long extension; and finally, only the smaller transcript has a potential TATA box.

interact with GTP-binding proteins. Less extensive similarity was found to the A5 protein of *Drosophila*, the OV16 antigen of *Onchocerca volvulus*, and DKA1 of yeast^{23–26}. The antennal protein A5 was isolated by differential screening, is expressed in subsets of olfactory hairs, and is possibly an odorant-binding protein. OV16 was identified as an antigen that is recognized by sera from individuals infected with the filarial parasite *O. volvulus*. The DKA1 gene has been shown to suppress certain *cdc25* mutations in yeast when introduced at high copy number, and may interact with the pathway involving RAS (GTP-binding protein), adenyl cyclase and membrane receptors of external stimuli. Mutants have only been generated for DKA1, and these have no phenotype.

The databases also revealed high similarity of *cen* to expressed sequence tags from *Arabidopsis* and rice. An *Arabidopsis* clone (129D7T7 from the *Arabidopsis* Biological Resource Center at Ohio State) was fully sequenced and showed 70% identity to *cen* at both the nucleotide and protein level, and preliminary data suggest that this is the *tfl* gene of *Arabidopsis* (manuscript in preparation). Our initial data suggest that the two rice clones (OSR29181A and OSS1946A, from the Rice Genome Research Project, Ibaraki, Japan) may have been derived from unspliced transcripts.

Expression of *cen* and *flo*

We have shown that *cen* normally acts to prevent *flo* expression in the apical meristem. To determine if this reflected its region of expression, we analysed the pattern of *cen* RNA by *in situ* hybridization (Fig. 3f). In mature, wild-type inflorescences the expression of *cen* was mostly limited to the region just below the apical meristematic dome, and was absent from the epidermal and outer cell layers. Traces of expression in the stem below the dome were limited to the region shown (Fig. 3f). Thus *cen* expression is subapical, and does not significantly overlap with the region of ectopic *flo* expression observed in the *cen* mutant, which occurs throughout the apical dome. These results indicate that *cen* acts at a distance to prevent *flo* expression in the apical dome, or that *cen* is below the level of direction in the apex. Expression of *cen* was detected in other tissues of wild-type plants by RT-PCR, but the significance of this has not yet been determined (data not shown).

As *cen* appears to regulate *flo* in the apical meristem, expression of *cen* should precede *flo*. To test this we looked at the timing of *cen* expression in relation to floral induction by collecting wild-type plants after 0, 1, 2, 4, 6 or 8 days of induction. RNA *in situ* hybridization analysis revealed that *cen* expression was not seen until 4 days, even though *flo* was induced in axillary meristems by 1–2 days. This raised the possibility that, in addition to acting upstream to *flo*, some aspects of *cen* might be downstream of *flo*. To test this, young inflorescence apices from *flo* mutant plants were probed for *cen* expression. Plants carrying the *flo*-640 allele, which produces non-functional *flo* RNA, showed no *cen* expression by RNA *in situ* hybridization (data not shown). Control sections probed with *flo* confirmed earlier analysis, that *flo*-640 RNA was present in young bract primordia and weakly in the dome of the apical meristem¹¹. It is therefore surprising that *flo* appears to be required for the expression of *cen* in the shoot apical meristem.

Models of inflorescence development

The architecture of the *Antirrhinum* inflorescence reflects the way in which *flo* and *cen* expression are coordinated. In *flo* mutants, flowers are replaced by indeterminate, continually branching shoots, whereas in *cen* mutants, the shoot is abruptly terminated by conversion to a flower. The gene expression studies suggest a model that accounts for the development of *cen* mutants and wild-type *Antirrhinum*.

In *cen* mutants, *flo* expression is activated within 1–2 days of floral induction (Fig. 6, *cen* day 2). Only the youngest initiating primordia are competent to form flowers, responding to the floral stimulus arriving at the apex. Older meristems remain vegetative,

either because they do not receive the inductive signal or they are not competent to respond. The apical meristem of *cen* mutants does not express *flo* up to about day 4 and continues to grow, generating axillary meristems on its flanks. The growth rate of the apex appears similar to wild type during this period, by the rate of node initiation and pattern of cell division, based on analysis of histone expression²⁷. However, by day 6, *flo* is expressed throughout the apical meristem, which is converted directly to a stage 2–3 floral meristem (Fig. 6). The ability of the apical meristem to be recruited to floral development at stages 2–3 contrasts with axillary meristems which cannot be recruited later than stage 0. This may be due to the inductive signal (or other factors) being spatially limited to the apical meristem.

There are two key features of this model that can account for additional aspects of the *cen* phenotype. First, the delayed activation of *flo* in the apical meristem allows several axillary meristems to be generated that give rise to flowers (usually 10). Second, the apical meristem is recruited at an advanced stage to become a stage 2–3 floral meristem. Expression of *flo* occurs throughout the meristem, and primordia that have already been initiated are recruited to form the sepals of the terminal flower. This suggests why the terminal flower is more advanced than axillary flowers, because at the time the apex is recruited the most advanced axillary meristem is only at stage 1–2. It further explains the spiral phyllotaxy of the sepals often observed in the terminal flower, as these could have been initiated by the apical inflorescence meristem just before it was recruited to form the terminal flower. Their variable number may simply reflect how many primordia were recruitable at the time the apical meristem was converted to floral development. The effects on organ number and symmetry of the terminal flower may also be due to direct conversion to a stage 2–3 floral meristem, without passing through stages 0–1. The earlier stages may be required for establishing the correct organ number and dorsal–ventral symmetry^{11,13}.

In wild-type apices, *flo* expression is activated in axillary meristems within 1–2 days of floral induction, as it is in *cen*

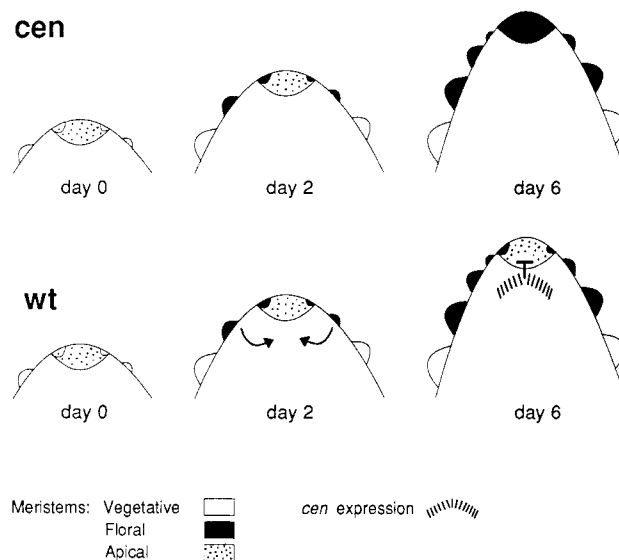


FIG. 6 Model for the early development of wild-type and *cen* mutant inflorescences of *Antirrhinum* upon floral induction. Diagrams show inflorescence apices on days 0, 2 and 6 after floral induction for the *cen* mutant and wild type. Axillary meristems develop from the flanks of the apical meristem. Floral meristems are indicated by black shading, apical meristems are dotted, and vegetative meristems have no shading. Bracts have been left out for clarity. In *cen* mutants, *flo* is activated in the apical meristem, but only after a delay. In wild type (*wt*), the arrows indicate signalling dependent on *flo* expression, leading to *cen* expression (line shading) in the sub-apical region of wild type. Activation of *cen* inhibits *flo* expression in the apical meristem.

mutants (Fig. 6, wild type day 2). During the next few days, further axillary meristems are generated by the apex, and these express *flo* and will give rise to flowers. No *cen* expression is observed in RNA *in situ* hybridization at this time (Fig. 6, day 2). However, by day 6, *cen* expression is found most strongly in the sub-apical region. Expression of *cen* appears to depend on the previous induction of *flo*, as RNA *in situ* analysis of young inflorescences of *flo* mutants indicated that they do not express *cen*. The possibility of a *flo*-independent pathway for the induction of *cen*, perhaps later in inflorescence development, awaits further analysis. Activation of *cen* inhibits apical *flo* expression, so that the meristem remains as an indeterminate inflorescence (Fig. 6, day 8). Expression of *flo* in axillary meristems is not inhibited by *cen*, either because *cen* action is spatially restricted or axillary meristems are less sensitive to *cen*. This model accounts for the observed ectopic accumulation of *flo* transcripts in the apex of *flo* mutants, as *cen* is not activated¹¹.

According to this model, *cen* and *flo* regulate each other non-autonomously. First, induction of *flo* activates *cen*, even though expression of *flo* is limited to axillary meristems and *cen* is activated in the sub-apical region. As these two regions do not overlap, this suggests that *flo* must signal across cells to activate *cen*. Second, in the *cen* mutant, *flo* is expressed ectopically throughout the apical dome, even though *cen* is largely expressed

below the dome and does not extend into the outer cell layers of the apex. Therefore *cen* represses *flo* in the wild-type apex by signalling across cells. How *flo* and *cen* act is not clear from their predicted proteins. The protein FLO may be a transcriptional activator and initiate a cascade of new gene activities that leads to cell-cell signalling⁹. In contrast, CEN may bind a lipid-like molecule or interact with GTP-binding proteins or the membrane. CEN may therefore mediate an inhibitory signal or prevent the flowering signal from reaching the apical meristem.

How might this model account for the evolution of indeterminate inflorescence architectures? One possibility is that the ancestral role of *cen* was to prevent flower initiation in vegetative meristems. The switch from vegetative to reproductive development would therefore have required repression of *cen*, leading to a determinate inflorescence. The indeterminate condition may have arisen by limiting *cen* activity to the apex after the reproductive switch, preventing the apical meristem from forming a flower. The ancestral role of *cen* in the control of flowering time may have been lost or may be masked by genetic redundancy in *Antirrhinum*. This predicts that ectopic expression of *cen* may prevent or reduce flowering. How widespread this mechanism is for the generation of different plant architectures will require the analysis of *cen* and *flo* homologues in other species, particularly those where *cen*-like mutants are available^{5,14-16}. □

Received 7 November 1995; accepted 15 January 1996.

- Weberling, F. *Morphology of Flowers and Inflorescences* (Cambridge Univ. Press, 1989).
- Coen, E. A. *Rev. Pl. Physiol. Pl. molec. Biol.* **42**, 241–279 (1991).
- Weigel, D. & Nilsson, O. *Nature* **377**, 495–550 (1995).
- Mandel, M. A. & Yanofsky, M. F. *Nature* **377**, 522–524 (1995).
- Coen, E. S. & Nugent, J. M. *Development* (suppl.) 107–116 (1994).
- Stebbins, G. L. *Flowering Plants, Evolution above the Species Level* (Harvard Univ. Press, MA, 1974).
- Kuckuck, H. & Schick, R. Z. *indukt. Abstamm. -u. Vererb. Lehre* **56**, 51–83 (1930).
- Stubbe, H. *Genetik und Zytologie von Antirrhinum L. sect Antirrhinum* (VEB Gustav Fischer, Jena, 1966).
- Coen, E. S. et al. *Cell* **63**, 1311–1322 (1990).
- Huijser, P. et al. *EMBO J.* **11**, 1239–1250 (1992).
- Carpenter, R. et al. *Pl. Cell* **7**, 2001–2011 (1995).
- Carpenter, R. & Coen, E. S. *Genes Dev.* **4**, 1483–1493 (1990).
- Coen, E. S. & Meyerowitz, E. M. *Nature* **353**, 31–37 (1991).
- Keeble, F., Pellew, C. & Jones, W. N. *New Phytol.* **9**, 68–77 (1910).
- Shannon, S. & Meeks-Wagner, D. R. *Pl. Cell* **3**, 877–892 (1991).
- Alvarez, J., Guli, C. L., Yu, X.-H. & Smyth, D. R. *Pl. J.* **2**, 103–116 (1992).
- Grandy, D. K. et al. *Molec. cell Endocr.* **4**, 1370–1376 (1990).
- Bucquoy, S., Jolles, P. & Schoentgen, F. *Eur. J. Biochem* **225**, 1203–1210 (1994).
- Schwarz-Sommer, Z. et al. *EMBO J.* **11**, 251–263 (1992).
- Bradley, D., Carpenter, R., Sommer, H., Hartley, N. & Coen, E. *Cell* **72**, 85–95 (1993).

- Coen, E. S., Robbins, T. P., Almeida, J., Hudson, A. & Carpenter, R. in *Mobile DNA* (eds Berg, D. E. & Howe, M. M.) 413–416 (American Society for Microbiology, Washington DC, 1989).
- Jameson, B. & Wolf, H. *Comput. Appl. Biosci.* **4**, 181–186 (1988).
- Pikielny, C. W., Hasan, G., Rouyer, F. & Rosbash, M. *Neuron* **12**, 35–49 (1994).
- Lobos, E. et al. *Molec. Biochem. Parasitol.* **39**, 135–146 (1990).
- Tripp, M. L., Bouchard, R. A. & Pinon, R. *Molec. Microbiol.* **3**, 1319–1327 (1989).
- Robinson, L. C. & Tatchell, K. *Molec. gen. Genet.* **230**, 241–250 (1991).
- Fobert, P. R., Coen, E. S., Murphy, G. J. P. & Doonan, J. H. *EMBO J.* **13**, 616–624 (1994).
- Williams, M. H. & Green, P. B. *Protoplasma* **147**, 77–79 (1988).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. molec. Biol.* **215**, 403–410 (1990).
- Luo, D., Coen, E. S., Doyle, S. & Carpenter, R. *Pl. J.* **1**, 59–69 (1991).
- Simon, R., Carpenter, R., Doyle, S. & Coen, E. *Cell* **78**, 99–107 (1994).
- Frohman, M. A., Dush, M. K. & Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8998–9002 (1988).
- Bradley, D., Vincent, C., Carpenter, R. & Coen, E. *Development* (in the press).
- Hammer, K., Knüpfer, S. & Knüpfer, H. *Kulturpflanze* **38**, 91–117 (1990).

ACKNOWLEDGEMENTS. We thank N. Hartley for sequencing of the genomic region of *cen*; P. Walker for care and maintenance of plants; Z. Ladbroke, G. Ingram and R. Waites for help with the plants in the summer; and P. Cubas, E. Schultz and R. Simon for comments and discussions. This work was supported by BBSRC PMB2 and Stem Cell Programmes, the EEC AMICA Programme and the Gatsby Foundation.

KNOW YOUR COPY RIGHTS

R E S P E C T O U R S

The publication you are reading is protected by copyright law.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

A new type of transient high-energy source in the direction of the Galactic Centre

C. Kouveliotou^{*†}, J. van Paradijs^{‡§}, G. J. Fishman[†],
M. S. Briggs[§], J. Kommers^{||}, B. A. Harmon[†],
C. A. Meegan[†] & W. H. G. Lewin^{||}

^{*} Universities Space Research Association, 4950 Corporate Drive, Suite 100, Huntsville, Alabama 35806, USA

[†] NASA Marshall Space Flight Center, ES-84, Huntsville, Alabama 35812, USA

[‡] Astronomical Institute "Anton Pannekoek" and Center for High-Energy Astrophysics, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

[§] Physics Department, University of Alabama in Huntsville, Huntsville, Alabama 35899, USA

^{||} Department of Physics, Center for Space Research, MIT 37-627, Cambridge, Massachusetts 02139, USA

SOURCES of high-energy (>20 keV) bursts fall into two distinct types: the non-repeating γ -ray bursters¹, several thousand of which have been detected but whose origin remains unknown, and the soft γ -ray repeaters (SGRs), of which there are only three². The SGRs are known to be associated with supernova remnants, suggesting that the burst events most probably originate from young neutron stars³. Here we report the detection of a third type of transient high-energy source. On 2 December 1995, we observed the onset of a sequence of hard X-ray bursts from a direction close to that of the Galactic Centre⁴. The interval between bursts was initially several minutes, but after two days, the burst rate had dropped to about one per hour and has been largely unchanged since then. More than 1,000 bursts have now been detected, with remarkably similar light curves and intensities; this behaviour is unprecedented among transient X-ray and γ -ray sources. We suggest that the origin of these bursts might be related to the spasmodic accretion of material onto a neutron star.

Visual inspection of the X-ray intensity records of the eight Large Area Detectors (LADs) of BATSE⁵ showed that bursts of

hard X-rays were emitted in rapid succession from a source in the general direction of the Galactic Centre beginning 1995 December 2 12:17 UT (Fig. 1). The bursts came at intervals of about 5 minutes; they lasted for typically 30 seconds and reached a peak intensity of ~ 300 counts s^{-1} in the 25–60 keV energy band. Spectral analysis of the events showed significant emission to ~ 60 keV in virtually all bursts; in $\sim 17\%$ of them, emission was detected to about 75 keV. The outbursts were also registered by the Spectroscopy Detectors⁵ which look into the same sky directions as the LADs and are sensitive down to ~ 10 keV. The bursts usually show a clear single peak; in many cases a tail of extended emission can be seen. The burst which we could unambiguously identify as the first in this sequence of rapid hard X-ray bursts occurred on December 2 08:22 UT. Between December 2 17:00 and 20:00 UT, the burst intervals clustered around (172 ± 15) s. After 20:00 UT the intervals became more erratic and longer. Figure 2 shows the behaviour of the burst interval evolution during the first 4 days of emission. Between 4 and 16 December the burst detection rate remained remarkably constant at an average of 18 events per day (about one-third of the bursts were missed because of data gaps).

We determined the location of the individual events from the relative strengths of the signals detected in different BATSE detectors, using an algorithm that includes a detailed description of the energy/angular detector response and the contribution from X-rays scattered off the Earth atmosphere⁶. Initially, the bursts were seen in two detectors only; to improve the constraints on the location, the Compton Gamma Ray Observatory (CGRO) was reoriented on December 8 00:57 UT, after which the events were registered by four detectors. The BATSE locations combined with (1) Earth occultation constraints of the source, (2) a galactic longitude constraint from observations with the Oriented Scintillation Spectroscopy Experiment (OSSE) on CGRO⁷, and (3) Interplanetary Network (IPN) triangulation results^{8,9}, produced an $\sim 24' \times 6'$ error box centred on right ascension 17 h 44 min 28 s, declination $-28^\circ 45.0'$ (equinox 2000.0). A 3-arcmin error box recently¹⁰ obtained with the X-Ray Timing Explorer (XTE) lies entirely within the IPN location. This error box does not contain known bright X-ray sources¹¹.

As part of a current project¹² we made a comprehensive search for events in the 20–60 keV energy range in archival BATSE data, covering the period 13 January 1993 to 24 December 1993 (345

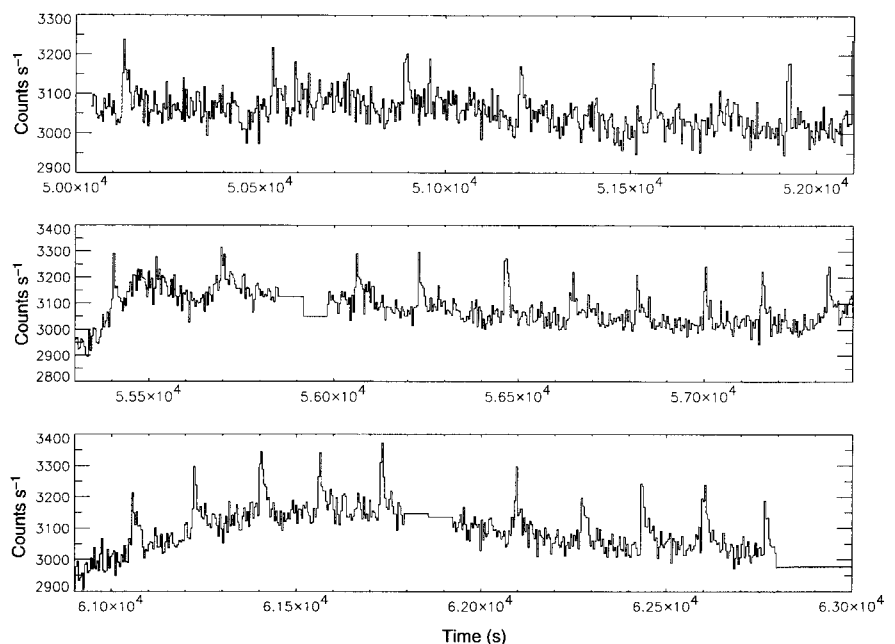


FIG. 1 The BATSE (25–60 keV) intensity records obtained during three consecutive orbits of CGRO on 2 December 1995 between 13:50 and 17:30 UT; the data are displayed with a time resolution of 4 s. (The time scale starts at 00:00 UT on 2 December 1995).

days). No activity which is unambiguously associated with the new transient was found.

In the following description of the properties of the hard X-ray bursts we limit ourselves to the 143 events detected during the first 4 days of burst activity, 134 of which have complete data. The durations of the bursts have been estimated from visual inspection of the X-ray intensity curves (25–60 keV, with 2-s time resolution). They range between ~ 6 and ~ 100 s (estimated accuracy ± 4 s), with the majority of the events between 10 and 30 s. Based on the count rates (and upper limits to them) in the CONT data² energy channels, we have made spectral fits, using several spectral models. Because of the limited statistics of the signals of individual events, we cannot make a significant distinction between a power-law fit (describing non-thermal emission) and thermal models, such as a black-body and thermal bremsstrahlung. The spectral parameters for the different bursts are rather similar; photon indices α range between -4.6 and -5.5 with an average value of -5.0 (average reduced χ^2 value is 1.73 (7 d.o.f.)). The typical kT values are ~ 5 keV for black-body fits and ~ 9 keV for thermal bremsstrahlung fits.

Using the power-law spectral fits we determined for each burst the (25–60 keV) peak flux, F_p , and the fluence S . The fits were made in order to convert count rates to fluxes and fluences. No particular physical meaning should be attached to the fit results, except as an indication of the hardness of the burst spectra. With few exceptions, the peak fluxes (25–60 keV) are confined to a fairly narrow interval. The average $\bar{F}_p = (2.8 \pm 0.7) \times 10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}$ (1σ). This may suggest that some limiting process is involved in the burst production. For a distance of 7.5 kpc (suggested by the small angular distance of the burst source to the Galactic Centre) the average (25–60 keV) peak luminosity is $2 \times 10^{38} \text{ erg s}^{-1}$. Although the bolometric luminosity depends on an uncertain extrapolation of the spectrum to lower energies, it is of interest to note that this value is of the same order as the Eddington luminosity of an object of one solar mass (assuming solar composition for the accreted material).

We have computed the integral hardness ratio per event, that is, the ratio of total counts in the 33–60 keV and 25–33 keV bands. The event hardness remains remarkably constant, with an average of 1.20 ± 0.24 . Unlike classical γ -ray bursters (GRBs), there is no dependence of hardness on duration¹³. Preliminary comparisons of the hardness ratios as measured for the first half of the counts in a burst (integrated over all energies) with that of the second half, find no evidence for spectral evolution during the bursts, similar to the bursts from soft γ -ray repeaters (SGRs).

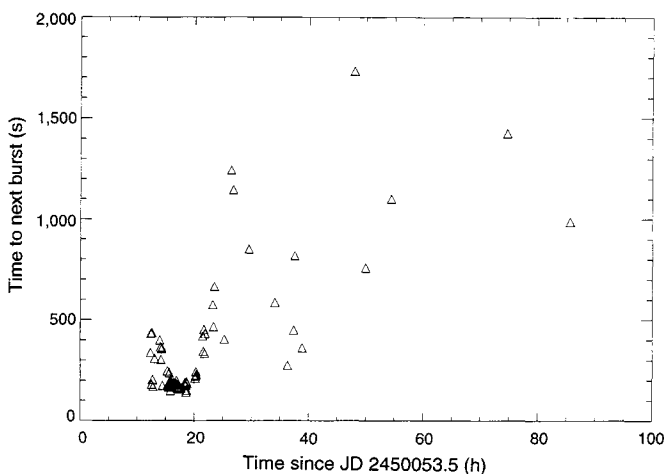


FIG. 2 Evolution of the intervals between bursts, from 14:00 UT on 2 December 1995 onwards. Note the very regular spacing of 172 s between bursts between 4 and 6 hours after the start of this sequence, and the subsequent increase and loss of regularity of the intervals.

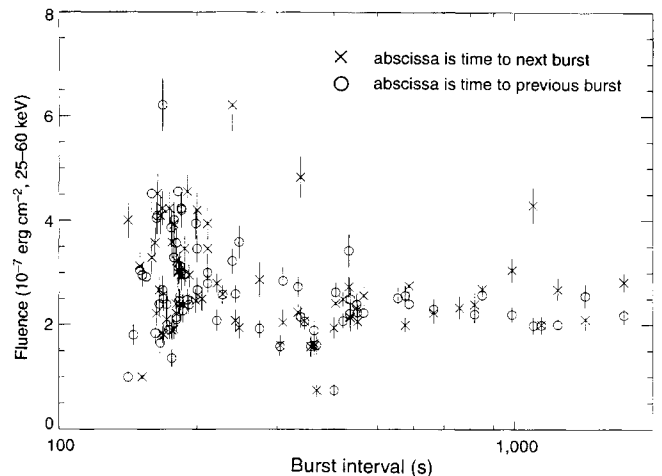


FIG. 3 Relation between the burst fluence and the time interval since the previous burst (crosses) and the time interval till the next burst (circles).

The burst fluences range between 1.7×10^{-7} and $6.8 \times 10^{-7} \text{ erg cm}^{-2}$; the average fluence $\bar{S} = (2.7 \pm 0.9) \times 10^{-7} \text{ erg cm}^{-2}$. In Fig 3 we show the relation between the burst fluence and the time interval before (Δt^-) and the interval after (Δt^+) the burst. The burst fluences do not show an obvious correlation with either time interval; for long intervals they appear to be almost constant.

The bursts we have detected are unlike those from any known source. They are very different from GRBs, which show no evidence of fast repetition and whose spectra are much harder (the bulk of their energy is above 100 keV; ref. 14). Their rapid recurrence is reminiscent of the rapidly repetitive type II bursts from the Rapid Burster¹⁵, whose nature still remains a mystery. However, the error box of the new burst source excludes the Rapid Burster.

With respect to their repetition and their spectra, the bursts we detected are similar to the events that have been observed from the three SGRs^{2,3,16}. However, they are very different in that they last several hundred times longer than SGR events (with the exception of two bursts from SGR0525 – 26 which lasted 3.5 and 8.5 s, respectively¹⁷). We can exclude the possibility that the hard X-ray bursts originate from the position¹⁸ of SGR1806 – 20 because one of them was observed when this source was occulted by the Earth.

To explain high-energy transients, three types of energy source have been proposed: (1) thermonuclear flashes in matter accreted on the surface of a neutron star; (2) unstable accretion onto a compact star; (3) internal energy sources of a neutron star (rotational and magnetic energy).

We can exclude the possibility that these bursts are caused by thermonuclear flashes on accreting neutron stars with low magnetic fields, as these sources produce the relatively well understood type I X-ray bursts¹⁵ whose spectra are described by black-body radiation, with temperature kT_{BB} generally below 2 keV, and which show distinct cooling during burst decay. Thermonuclear flashes on the surface of a strongly magnetized neutron star (if they occur at all) have been proposed as a model for GRBs¹⁹ and may have much harder spectra than type I bursts. However, such flashes should have very long recurrence times.

Independent of detailed thermonuclear flash models, during the period of rapid burst repetition the required replenishment of burnt fuel (through accretion onto the neutron star) would lead to an accretion flux of the order of several times $10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}$, that is, one would expect a very bright X-ray source. On the basis of the BATSE occultation data, which provides an upper limit of $\sim 10^{-9} \text{ erg cm}^{-2} \text{ s}^{-1}$ to the persistent flux, we can exclude the possibility that this persistent accretion power is emitted as X-rays with energies in excess of ~ 10 keV. For the ratio of persistent flux

to time-averaged burst flux (α , ref. 15), as measured in the BATSE energy range, we find a 3σ upper limit of ≤ 4 . These results virtually rule out the possibility that the bursts are caused by thermonuclear flashes on neutron stars.

We consider accretion instabilities a viable model for the bursts from the Galactic Centre region. If this model is correct, it implies that the source of hard X-ray bursts is part of a binary system. In that case, it is likely that the matter is accreted onto the compact object in spasmodic fashion after storage in an accretion disk. Disk storage has been suggested as a model for the type II bursts from the Rapid Burster²⁰; however, the hard X-ray bursts that we observed differ from the type II bursts of the Rapid Burster by their much harder spectra, and by not following the relaxation oscillator recurrence pattern shown by that source. This does not, however, exclude the possibility that the bursts reported here are caused by some accretion instability. In particular, the almost regular intervals between the bursts are much more reminiscent of the Rapid Burster than of the SGR sources, pointing to a triggering mechanism similar to the one for type II bursts rather than to SGR bursts.

According to the SGR model of Thomson and Duncan²¹, hard X-ray events may originate from strongly magnetized neutron stars (magnetars) through the release of magnetic field energy, which leads to the formation of a pair plasma confined by the magnetosphere. The duration of the events is determined by the diffusion time of the plasma, which is proportional to the inverse square of the magnetic field strength. Although we favour an accretion instability origin for the bursts, we believe that magnetar models cannot be excluded.

If the bursts reported here are due to accretion instabilities, which is likely, they are by definition¹⁵ type II bursts. The question then remains why are they so different from those of the Rapid Burster. The answer could lie with the compact object which may

be a neutron star with a much higher magnetic field than the Rapid Burster.

Note added in proof: Since this Letter was submitted, a persistent source, GRO J1744–28, has been found²² and later identified as a binary X-ray pulsar^{23,24}. The burster and the pulsar are the same object²⁵. □

Received 3 January; accepted 5 February 1996.

1. Fishman, G. J. & Meegan, C. A. *Rev. Astr. Astrophys.* **33**, 415–458 (1995).
2. Norris, J. P., Hertz, P., Wood, K. & Kouveliotou, C. *Astrophys. J.* **366**, 240–252 (1991).
3. Kuikarni, S. R., Frail, D. A., Kassim, N. E., Murakami, T. & Vasisth, G. *Nature* **368**, 129–131 (1994).
4. Fishman, G. J. *et al.* *IAU Circ. No.* 6272 (1995).
5. Fishman, G. J. *et al.* in *Proc. Gamma Ray Observatory Science Workshop* (ed. Johnson, W. N.) 39–50 (NASA, Goddard Space Flight Center, Greenbelt, MD, 1989).
6. Pendleton, G. *et al.* *Astrophys. J.* (in the press).
7. Kurfess, J. D. *et al.* *IAU Circ. No.* 6278 (1995).
8. Hurley, K. *et al.* *IAU Circ. No.* 6276 (1995).
9. Hurley, K. *et al.* *IAU Circ. No.* 6286 (1996).
10. Swank, J. *et al.* *IAU Circ. No.* 6291 (1996).
11. Van Paradijs, J. in *X-ray Binaries* (eds Lewin, W. H. G. *et al.*) 536–577 (Cambridge Univ. Press, 1995).
12. Kommers, J. *et al.* in *Proc. 3rd Huntsville Gamma Ray Burst Workshop* (eds Kouveliotou, C., Briggs, M. S. & Fishman, G. J.) (in the press).
13. Kouveliotou, C. *et al.* *Astrophys. J.* **413**, L101–L104 (1993).
14. Band, D. *et al.* *Astrophys. J.* **413**, 281–292 (1993).
15. Lewin, W. H. G., Van Paradijs, J. & Taam, R. E. in *X-ray Binaries* (ed. Lewin, W. H. G. *et al.*) 175–232 (Cambridge Univ. Press, 1995).
16. Kouveliotou, C. *et al.* *Nature* **362**, 728–730 (1993).
17. Mazets, E. P. *et al.* *Astrophys. Space Sci.* **80**, 3–143 (1981).
18. Van Kerkwijk, M. H. *et al.* *Astrophys. J.* **440**, L61–L64 (1995).
19. Woosley, S. E. & Wallace, R. K. *Astrophys. J.* **258**, 716–732 (1982).
20. Spruit, H. C. & Taam, R. E. *Astrophys. J.* **402**, 593–604 (1993).
21. Thomson, C. & Duncan, R. G. *Astrophys. J.* (in the press).
22. Paciesas, W. S., Harmon, B. A., Fishman, G. J., Zhang, S. N. & Robinson, C. R. *IAU Circ. No.* 6284 (1996).
23. Finger, M. H., Wilson, R. B., Harmon, B. A., Hagedorn, K. & Prince, T. A. *IAU Circ. No.* 6285 (1996).
24. Finger, M. H., Wilson, R. B. & van Paradijs, J. *IAU Circ. No.* 6286 (1996).
25. Kouveliotou, C. *et al.* *IAU Circ. No.* 6286 (1996).

ACKNOWLEDGEMENTS. J.V.P., C.K. and W.H.G.L. were supported by NASA; J.M.K. was supported by a National Science Foundation graduate research fellowship.

Re-entry and ablation of cometary dust in the impact plumes of Shoemaker–Levy 9

A. Fitzsimmons*, P. J. Andrews†, R. Catchpole‡, J. E. Little*, N. Walton‡ & I. P. Williams§

* Department of Pure and Applied Physics, Queens University of Belfast, Belfast BT7 1NN, UK

† Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ, UK

‡ Royal Greenwich Observatory, Apartado 321, 38780 Santa Cruz de La Palma, Spain

§ Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

COLLISIONS between small bodies (such as asteroids and comets) and the terrestrial planets are known to throw ejecta far beyond the point of impact. But until the fragments of comet Shoemaker–Levy 9 hit Jupiter in July 1994 (refs 1, 2), there had been no opportunity to study the effects of such collisions on gas-giant planets. Here we present optical spectra obtained during the collision of fragments L and Q₁ with Jupiter. We observed emission lines from sodium, magnesium, calcium, iron, manganese and chromium as the ejecta plume fell back onto Jupiter's atmosphere. All of these elements are expected to occur only very deep in Jupiter's atmosphere—considerably below the depth to which the fragments penetrated—, suggesting that the material responsible for the emissions originated almost entirely in the comet itself. The initial phase of emission is associated with heating of the impact ejecta as it fell back onto the planet. A second, later phase of emission was also observed, which we

associate with grains of silicate dust that condensed within the cooling fireball and subsequently re-entered, meteor-like, into Jupiter's atmosphere.

Spectra were obtained using the Intermediate Dispersion Spectrograph on the 2.5-m Isaac Newton Telescope on La Palma, together with a blue-sensitive Tektronix 1,024 × 1,024 pixel CCD (charge-coupled device) as detector. On the night of 19 July 1994, the R300V grating was used to give a wavelength coverage of 3,600–6,600 Å. The slit width of 1.5 arcsec on the sky resulted in a resolution of 6.2 Å (full width at half maximum) at this wavelength. The slit was placed tangential to the limb of Jupiter at the expected impact latitude of -43° , and exposures of 3 seconds duration were taken approximately every 2 minutes. Spectra were obtained continuously from 22:01 to 22:42 UT, covering the expected time of impact of fragment L.

The time of impact has been given as 22:16:48 UT from observations obtained by the Galileo spacecraft³. We first observed weak emission at 22:26 UT due to the unresolved Na D lines at 5,890 Å. This emission saturated the detector for the next three exposures, allowing only lower limits on the flux to be determined, and then diminished. Emission due to other species (primarily Mg I, Ca I and Fe I) first appeared at 22:31 UT, reached a broad maximum between 22:33 and 22:38 UT, and was rapidly fading by the time of our last observation. The evolution of our spectra during the observation period is shown in Fig. 1a. Given the difficulty of aligning the slit on the observed impact site during these observations (owing to the strong interference of scattered light on the jovian limb), we believe that the small irregularities in the emission lightcurve (that is, at 18 minutes after impact) are probably caused by guiding errors and that the entire evolution was in reality more smooth. As all models of the atmospheric structure place the elements observed at depths below 10^5 bar, far deeper than the expected maximum penetration depth of the bolides, the

observed material must have originated entirely within the comet fragments.

In Fig. 1b we show the lightcurve at 9,069 Å of the observed ejecta plume as it rose above the limb of Jupiter and then fell back onto the planet⁴. The strongest Na I emission was contemporaneous with the fallback of the ejecta plume and the peak of emission in the infrared⁵, a signature of the heating of the jovian atmosphere by the re-entering ejecta. We have modelled the Na I emission assuming that it arises from thermal excitation of the ejecta within the heated jovian stratosphere. From imaging by the Hubble Space Telescope¹ and La Palma⁴, the impact region was assumed to be an edge-on disk of radius $\sim 5 \times 10^3$ km and thickness ~ 10 km. Although the observed ejecta fallout pattern is more complex¹, a more accurate representation of the ejecta impact region would not affect our results significantly. Combining this with the observed Na fluxes for the strongest (overexposed) emission, we derive a lower limit to the temperature of $T \geq 790$ K and a mean column density of $N(\text{Na}) > 10^{12} \text{ cm}^{-2}$. Assuming that the original bolide possessed near-solar abundances, this implies a total ejected mass of $> 5 \times 10^8$ g, equivalent to a cometary nucleus of radius > 5 m if its density $\rho = 1 \text{ g cm}^{-3}$. Although this is only a lower limit, this mass estimate supports the

TABLE 1 Atomic emission lines observed during the impact of fragment L

Species	Wavelength (Å)	E_l (cm ⁻¹)	E_u (cm ⁻¹)
Ca I	6,573	0	15,210
Na I*	5,890	0	16,973
Na I*	5,896	0	16,956
Mn I*	4,032	0	24,802
Mn I*	4,034	0	24,788
Mn I*	4,036	0	24,779
Mg I	4,571	0	21,870
Mn I	5,395	0	18,470
Mn I	5,433	0	18,309
Fe I	4,376	0	22,846
Fe I	4,427	416	22,997
Fe I	4,463	704	23,111
Fe I	5,110	0	19,562
Fe I*	5,169	416	19,757
Fe I*	5,166	0	19,351
Fe I	5,204	704	19,912
Fe I	5,225	888	20,020
Fe I	5,250	978	20,020
Cr I	4,254	0	23,390
Cr I	4,274	0	23,309

E_l and E_u are the energies above the ground state of the lower and upper energy levels for the observed transitions.

* Blended in the observed spectrum.

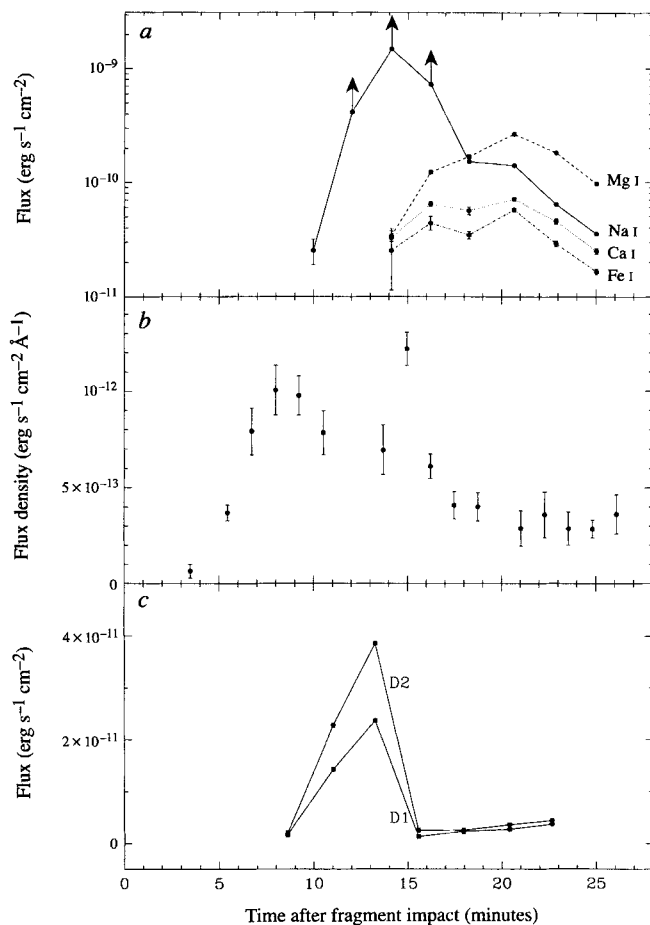


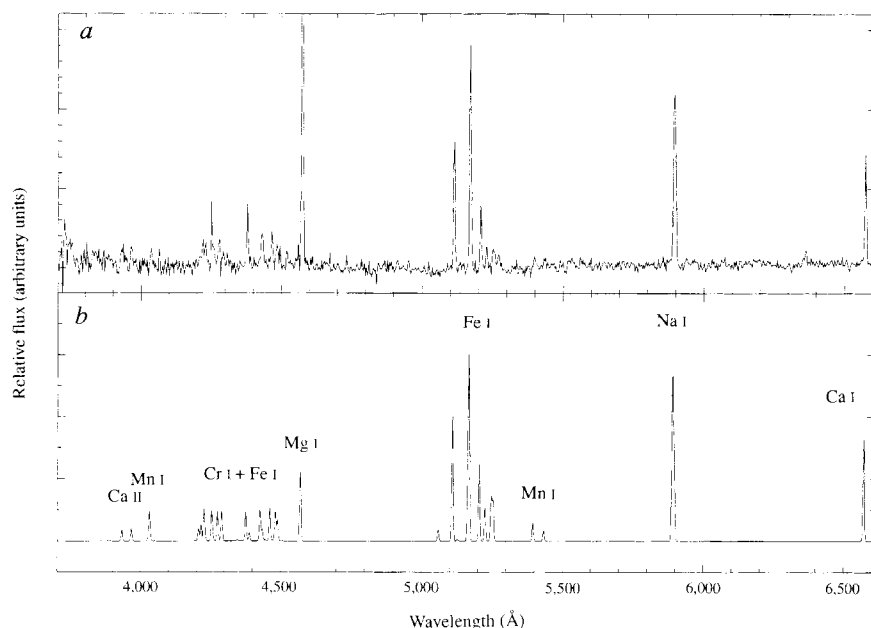
FIG. 1 a, The observed emission line strengths for strong lines (Na I 5,889 Å (D2) and 5,896 Å (D1; blended), Ca I 6,572 Å, Fe I 5,110 Å, Mg I 4,572 Å) after the impact on Jupiter of fragment L on 19 July 1994. Error bars show the uncertainty due to the line measurement. Uncertainty in the absolute flux calibration is estimated at 10%. Large black arrows indicate that fluxes are lower limits due to detector saturation. b, The lightcurve of the ejecta measured simultaneously at 9,069 Å using the 1-m Jacobus Kapteyn Telescope on La Palma⁴. From comparison with a the spectroscopic emission can clearly be seen to be associated with the splashback phase of the ejecta plume. c, The observed strengths of the Na I D2 and D1 lines following the impact of fragment Q₁ on 20 July 1994. Errors as in a.

findings of theoretical models which predict that only a small fraction of the impactor mass went into the plume itself⁶, and implies that the best estimates of impactor size may come from total abundance measurements of the impact sites⁷.

The emission due to other metals such as Ca I, Fe I and Mg I seems to be connected with the latter part of the re-entry phase. Up to now the mechanism causing these emissions has not been established. Roos-Serote *et al.*⁸ observed the same impact at higher spectral, but poorer temporal, resolution, and concluded the fluorescence by scattering of sunlight alone could not reproduce the observed line ratios if the lines were optically thin. In particular, we have calculated the Ca I 6,573 Å to Ca I 4,226 Å line ratio in the case of fluorescence at intermediate and large optical depths, and come to the same conclusion for all column densities. This is because optically thin emission would result in a line ratio of $\sim 10^{-5}$, while at high optical depths all solar photons would be scattered, and the line ratio would be approximately the ratio of the incident solar flux at the two wavelengths.

Instead, by comparing the spectral evolution with that of the visible ejecta, we interpret the appearance of the majority of emission lines beginning at 22:31 UT as being caused by the meteoritic ablation of re-entering dust particles. The ablated metal atoms then produce the observed emission through simple thermal excitation in the infalling meteors. To interpret our data we used a model similar to those used previously for meteor spectra⁹, in which the emitting region is at a constant temperature and in local thermodynamic equilibrium. Although neither of these assumptions is likely to be correct, it does allow an initial interpretation of the phenomena observed. Atomic data for the observed species were gathered from various sources¹⁰⁻¹². We also initially assume an electron density of $n_e = 10^3 \text{ cm}^{-3}$, as this is the minimum observed in the jovian stratosphere¹³. The atomic emission lines that we observed are shown in Table 1. These data indicate that only near-ground-state emissions were observed by us, implying a low temperature for the ablating dust. Other observers have observed emissions from metastable levels above the ground state. Such levels might be populated through collisional excitation in the meteor head, which is not accounted for in our model. Another possibility is that such lines are due to natural decay from high-level metastable states that were populated in the initial fireball¹⁴.

FIG. 2 *a*, The observed emission spectrum of the impact site of fragment L at 22:37 UT. *b*, Theoretical spectrum calculated for $T = 1,700$ K, $n_e = 10^3 \text{ cm}^{-3}$, $N(\text{Fe}) = 5 \times 10^{16}$ and relative number densities for other elements given in the text. The apparent emission at 4,260 Å is in reality a noise spike. Optical depths at the line centre of Ca I 4,226 Å and Ca I 6,572 Å are 2,400 and 0.1 respectively, explaining the observed line ratios. Similarly, the saturation of the Mg I 4,572 Å emission (line centre optical depth, 130) allows only a lower limit to be placed on the relative abundance of Mg.



The relative strengths of the Fe I multiplet 1 and multiplet 2 emission demonstrate a rapid rise in the excitation temperature, climbing from $T = 1,500 \pm 100$ K at 22:33 UT to $1,800 \pm 100$ K at 22:37 UT, and column densities in the emitting region increasing from $N(\text{Fe}) = 10^{16} \text{ cm}^{-2}$ to 10^{17} cm^{-2} over the same period. This is what would be expected from ballistically re-striking grains, as those arriving at later times would be re-entering at higher velocities and would be heated to higher temperatures. These low temperatures then explain why emission from the metastable levels of Ca I and Mg I are so strong, as they are the only levels at a low enough energy in these atoms to become significantly populated.

In Fig. 2 we compare the spectrum with the highest signal-to-noise ratio (obtained at 22:37 UT) with a theoretical spectrum calculated for number densities (relative to Fe) of 10 for Mg, 10^{-3} for Na, 2×10^{-2} for Ca, 3×10^{-3} for Mn and 10^{-3} for Cr. This is clearly a good representation of the observed emission pattern. The error of ± 100 K in T corresponds to a factor of ± 10 uncertainty in the Na column density and a factor of 2–4 in that of Ca. The largest uncertainty in these abundances arises from the uncertainty in the degree of ionization and hence n_e , which in turn most seriously affects the derived Na and Ca column densities. If n_e was increased to $\sim 10^{13} \text{ cm}^{-3}$ as observed in terrestrial meteors⁹, then Ca and Na would be further depleted by factors of 1.5 and 50 respectively. The large underabundance of Na in the grains implies that little went into grain formation, and that the above calculations of total mass ejected remain valid.

Of particular note is the strong Mg I emission which cannot be reproduced in our model with a single emitting grain population, as the line quickly becomes saturated at even modest enhancements. Although some collisional ionization of the observed species may be present, it cannot be a prime factor in the observed abundances as the low temperatures would allow immediate recombination, and such transitions have not been observed. As the observed fluxes show that the emission from the in-falling meteors filled only a small fraction of our aperture, this implies that the particles must have consisted of at least two populations, one of which contributes significant amounts of Mg.

Considering these points, dominance of Mg and Fe with respect to other elements must be a reflection of the grain abundances themselves, and points to a silicate composition for the dust, dominated by minerals such as MgSiO_3 (enstatite) and FeSiO_3 (ferrosilite). The condensation temperatures of such minerals is $\sim 1,500$ K, in good agreement with the temperatures found

above, and supporting the ablating-grain hypothesis. This also offers an explanation as to why Fe I and Mg I was not observed before 22:31 UT even though grain re-entry was occurring at this time, as extrapolating from the Fe I lines the temperature was $< 1,500$ K and therefore significant ablation could not occur.

Observational evidence for silicates and their atomic constituents in other impacts has been reported^{15,16}. Theoretical calculations show that silicates may form in the early stages of plume development^{2,17}, with possible grain sizes easily reaching ~ 1 – $100 \mu\text{m}$. From calculations (D. Hughes, personal communication) and time of flight considerations, the emission must come from ablation of grains with masses $\geq 10^{-3}$ g with radii $> 500 \mu\text{m}$ for a mean density $\rho \leq 2.5 \text{ g cm}^{-3}$. Such large grains imply that the initial fireball must have been relatively compact and dense, presumably due to the (inferred) large mass of the impactor. Whatever their origin, the observed emission demonstrates that this large grain population was still re-entering the jovian stratosphere some 25 minutes after the initial impact of fragment L of Shoemaker–Levy 9.

On the 20 July we also obtained spectra of the impact of fragment Q₁. The telescope and spectrograph were the same, but the R1200Y grating was used to observe the wavelength range 5,620–6,120 Å at a resolution of 1.3 Å. The only emission lines observed were the Na D lines at 5,889 and 5,896 Å. The spectral evolution is shown in Fig. 1c and was qualitatively similar to that of the L-impact the previous night, in that a pronounced maximum in the strength of the Na emission occurs ~ 14 minutes after the bolide entry observed by Galileo at 20:13:52 UT³. The fluxes for the Na emission associated with the Q₁ impact imply that it involved significantly less material than that of fragment L. The Na doublet was resolved in these observations with a D_2/D_1 flux ratio of 1.61 ± 0.02 when the emission was strongest. This line ratio strongly constrains the column density in the emitting region to $(1\text{--}3) \times 10^{11} \text{ cm}^{-2}$ for temperatures up to 2,000 K. For an emitting area equivalent to that of the L-impact assumed above, the temperature of this region was $T = 740$ K, and the total amount of Na detected by our spectrograph was $\sim 1 \times 10^8$ g, equivalent to a 3-m-radius nucleus with a density of 1 g cm^{-3} . □

Received 23 October 1995; accepted 24 January 1996.

1. Hammel, H.B. *et al.* *Science* **267**, 1288–1296 (1995).
2. West, R.A. *et al.* *Science* **267**, 1296–1301 (1995).
3. Martin, T.Z., Orton, G.S., Travis, C.D., Tampari, L.K. & Llaypool, I. *Science* **268**, 1875–1879 (1995).
4. Fitzsimmons, A. *et al.* *Mon. Not. R. astr. Soc.* **278**, 781–786 (1996).

5. Lagage, P. O. et al. in *European SL-9/Jupiter Workshop* (eds West, R. & Böhmhardt, H.) 147–152 (ESO, Garching, 1995).
6. Takata, T., O'Keefe, J. D., Ahrens, T. J. & Orton, G. S. *Icarus* **109**, 3–19 (1994).
7. Lellouch, E. et al. *Nature* **373**, 592–595 (1995).
8. Roos-Serote, M. et al. *Geophys. Res. Lett.* **22**, 1621–1624 (1995).
9. Borovicka, J. *Astr. Astrophys.* **279**, 627–645 (1993).
10. Wiese, W. L. (ed.) in *Spectroscopic Data for Iron* (ORNL-6089/V4, Oak Ridge National Lab., Tennessee, 1985).
11. Wiese, W. L., Smith, M. W. & Miles, B. M. in *Atomic Transition Probabilities* (NSRDS-NBS 22, National Bureau of Standards, Washington DC, 1969).
12. Kurucz, R. L. *Atomic Data Files* (Harvard-Smithsonian Center for Astrophysics, World-Wide Web <http://cfa-www.harvard.edu/amp/data/stats/hvrvz.html> 1995).
13. Atreya, S. K. & Donahue, T. M. in *Jupiter* (ed. Gehrels, T.) 304–318 (Univ. Arizona Press, Tucson, 1976).
14. Carpenter, K. G. et al. *Bull. Am. astr. Soc.* **27**, 1118 (1995).
15. Nicholson, P. D. et al. *Geophys. Res. Lett.* **22**, 1621–1624 (1995).
16. Noll, K. S. et al. *Science* **267**, 1307–1313 (1995).
17. Hasegawa, H., Takeuchi, S. & Watanabe, J. in *European SL-9/Jupiter Workshop* (eds West, R. & Böhmhardt, H.) 279–285 (ESO, Garching, 1995).

ACKNOWLEDGEMENTS. We thank the staff on La Palma for their support during the observing period. The Isaac Newton telescope is operated on the island of La Palma by the Royal Greenwich Observatory in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. J.E.L. was supported by PPARC.

Scaling behaviour in the growth of companies

Michael H. R. Stanley*, Luís A. N. Amaral*,
Sergey V. Buldyrev*, Shlomo Havlin*†,†,
Heiko Leschhorn*, Philipp Maass*,
Michael A. Salinger‡ & H. Eugene Stanley*

* Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA

† Minerva Center and Department of Physics, Bar-Ilan University, Ramat Gan, Israel

‡ Department of Finance and Economics, School of Management, Boston University, Boston, Massachusetts 02215, USA

A SUCCESSFUL theory of corporate growth should include both the external and internal factors that affect the growth of a company^{1–18}. Whereas traditional models emphasize production-related influences such as investment in physical capital and in research and development¹⁸, recent models^{10–20} recognize the equal importance of organizational infrastructure. Unfortunately, no exhaustive empirical account of the growth of companies exists by which these models can be tested. Here we present a broad, phenomenological picture of the dependence of growth on company size, derived from data for all publicly traded US manufacturing companies between 1975 and 1991. We find that, for firms with similar sales, the distribution of annual (logarithmic) growth rates has an exponential form; the spread in the distribution of rates decreases with increasing sales as a power law over seven orders of magnitude. A model wherein the probability of a company's growth depends on its past as well as present sales accounts for the former observation. As the latter observation applies to companies that manufacture products of all kinds, organizational structures common to all firms might well be stronger determinants of growth than production-related factors, which differ for companies producing different goods.

The simplest model for corporate growth was proposed by Gibrat¹. Its basic assumptions are that the rate of company growth is (1) independent of company size (law of proportionate effect), and (2) uncorrelated in time. These assumptions can be formalized by the following random multiplicative process: $S_{t+\Delta t} = S_t(1 + \epsilon_t)$, where $S_{t+\Delta t}$ and S_t are the sales of the company at time $t + \Delta t$ and t respectively, and ϵ_t is an uncorrelated random number with mean close to zero and standard deviation much smaller than one. Hence $\log S_t$ follows a simple random walk so that firm sizes are log-normally distributed. Also, for sufficiently large time intervals $T \gg \Delta t$, the growth rates S_{t+T}/S_t are log-normally distributed.

Although it is known that Gibrat's assumptions are rejected

empirically, many theoretical and empirical analyses still use the Gibrat model as a benchmark, for lack of a better alternative^{21–26}. To achieve a more realistic characterization of company dynamics, we analyse the statistical properties of the growth rates.

We studied all US manufacturing publicly traded companies within the years 1975–91. The data were taken from the Compustat database and all values for sales have been adjusted to 1987 dollars by the GNP price deflator. (Compustat contains financial information that firms traded on United States securities exchanges must file with the US Securities and Exchange Commission.) We define a firm's annual growth rate as $R \equiv S_1/S_0$, where S_0 and S_1 are its sales in two consecutive years.

It is customary to study company growth on logarithmic scales, so we define $r \equiv \ln(S_1/S_0)$ and $s_0 \equiv \ln S_0$ and calculate the conditional distribution $p(r | s_0)$ of growth rates r with a given initial sales value s_0 .

The distribution $p(r | s_0)$ of the growth rates from 1990 to 1991 is shown in Fig. 1a for two different values of initial sales. Remarkably, both curves display a simple 'tent-shaped' form. The distribution is not gaussian—as expected from the Gibrat model—but rather is exponential,

$$p(r | s_0) = \frac{1}{\sqrt{2}\sigma(s_0)} \exp\left(-\frac{\sqrt{2}|r - \bar{r}(s_0)|}{\sigma(s_0)}\right) \quad (1)$$

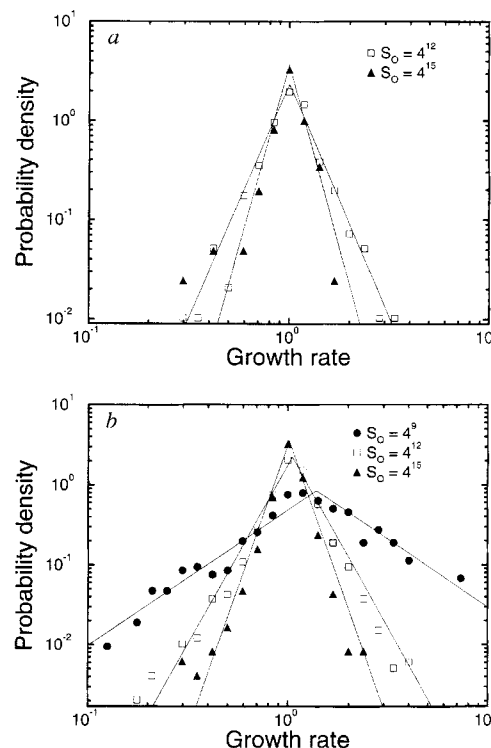


FIG. 1 a, Probability density $p(r | s_0)$ of the growth rate $r \equiv \ln(S_1/S_0)$ from year 1990 to 1991 for all publicly traded US manufacturing firms in the 1994 Compustat database with standard industrial classification index of 2000–3999. We examine 1991 because between 1992 and 1994 there are several companies with zero sales that either have gone out of business or are 'new technology' companies (developing new products). We show the data for two different bins of initial sales (with sizes increasing by powers of 4): $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). Within each sales bin, each firm has a different value of R , so the abscissa value is obtained by binning these R values. The solid lines are fits to equation (1) (in the text) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from the data. b, Probability density $p(r | s_0)$ of the annual growth rate, for three different bins of initial sales: $4^{8.5} < S_0 < 4^{9.5}$ (circles), $4^{11.5} < S_0 < 4^{12.5}$ (squares) and $4^{14.5} < S_0 < 4^{15.5}$ (triangles). The data were averaged over all 16 one-year periods between 1975 and 1991. The solid lines are fits to equation (1) using the mean $\bar{r}(s_0)$ and standard deviation $\sigma(s_0)$ calculated from all data.

The straight lines shown in Fig. 1a are calculated from the average growth rate $\bar{r}(s_0)$ and the standard deviation $\sigma(s_0)$ obtained by fitting the data set to equation (1).

We also find that the data for each of the 16 annual intervals from the period 1975–91 fit well to equation (1), with only small variations in the parameters $\bar{r}(s_0)$ and $\sigma(s_0)$. To improve the statistics, we therefore calculate the new distribution by averaging all the data from the 16 annual intervals in the database. As shown in Fig. 1b, the data now scatter much less and the shape is well described by equation (1). For this reason, we have also included in the figure data for 'volatile' cases, corresponding to sales of only about 2.6×10^5 dollars.

As is apparent from Fig. 1b, $\sigma(s_0)$ decreases with increasing s_0 . We find $\sigma(s_0)$ is well approximated over more than seven orders of magnitude—from sales of less than 10^4 dollars up to sales of more than 10^{11} dollars—by the law

$$\sigma(s_0) = a \exp(-\beta s_0) = a S_0^{-\beta} \quad (2)$$

where $a \approx 6.66$ and $\beta = 0.15 \pm 0.03$ (Fig. 2).

We performed a parallel analysis for the number of employees, and the corresponding standard deviation is shown in Fig. 2. The data are linear over roughly five orders of magnitude, from firms with only 10 employees to firms with almost 10^6 employees. The slope $\beta = 0.16 \pm 0.03$ is the same, within error bars, as that found for sales.

We find that equations (1) and (2) accurately describe three additional indicators of company growth; (1) cost of goods sold (with exponent $\beta = 0.16 \pm 0.03$), (2) assets ($\beta = 0.17 \pm 0.04$), and (3) property, plant and equipment ($\beta = 0.18 \pm 0.03$).

What is remarkable about equations (1) and (2) is that they govern the growth rates of a diverse set of firms. They range not only in their size but also in what they manufacture. The conventional economic theory of the firm is based on production technology, which varies from product to product. Conventional theory does not suggest that the processes governing the growth rate of car companies should be the same as those governing, for example, pharmaceutical or paper firms. Indeed, our findings are reminiscent of the concept of universality found in statistical physics, where different systems can be characterized by the same fundamental laws, independent of 'microscopic' details.

In statistical physics, scaling phenomena of the sort that we have uncovered in the sales and employee distribution functions are sometimes represented graphically by plotting a suitably 'scaled' dependent variable as a function of a suitably 'scaled' independent variable. If scaling holds, then the data for a wide range of parameter values are said to 'collapse' upon a single curve. To test the present data for such data collapse, we plot (Fig. 3) the

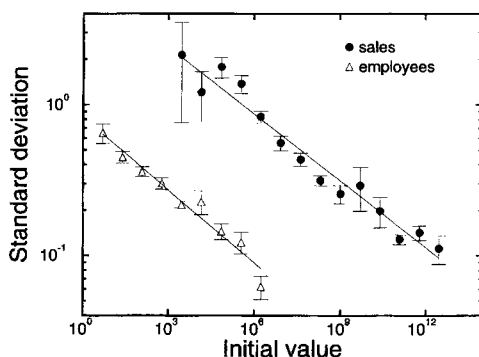


FIG. 2 Standard deviation of the one-year growth rates of the sales (circles) and of the one-year growth rates of the number of employees (triangles) as a function of the initial values. The solid lines are least-square fits to the data with slopes $\beta = 0.15 \pm 0.03$ for the sales and $\beta = 0.16 \pm 0.03$ for the number of employees. We also show error bars of one standard deviation about each data point. These error bars appear asymmetric as the ordinate is a log scale.

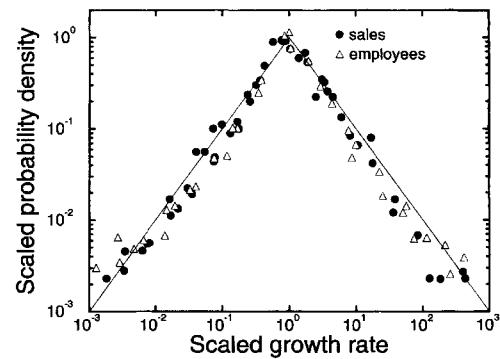


FIG. 3 Scaled probability density $p_{\text{scal}} \equiv 2^{1/2} \sigma(s_0) p(r | s_0)$ as a function of the scaled growth rate $r_{\text{scal}} \equiv 2^{1/2} [r - \bar{r}(s_0)] / \sigma(s_0)$ of sales (circles). The values were rescaled using the measured values of $\bar{r}(s_0)$ and $\sigma(s_0)$. Also we show (triangles) the analogous scaled quantities for the number of employees. All the data collapse upon the universal curve $p_{\text{scal}} = \exp(-|r_{\text{scal}}|)$ (solid line) as predicted by equations (1) and (2).

scaled probability density $p_{\text{scal}} \equiv \sqrt{2} \sigma(s_0) p(r | s_0)$ as a function of the scaled growth rates of both sales and employees $r_{\text{scal}} \equiv \sqrt{2} [r - \bar{r}(s_0)] / \sigma(s_0)$. The data collapse upon the single curve $p_{\text{scal}} = \exp(-|r_{\text{scal}}|)$. Our results for (1) cost of goods sold, (2) assets, and (3) property, plant and equipment are equally consistent with such scaling.

The Gibrat model, which yields a log-normal distribution of the growth rates for sufficiently long time intervals, fails to explain the observed distribution of annual growth rates (even for intervals as long as 5 years, we find $p(r | s_0)$ does not obey a normal distribution.) There is, however, a simple dynamic process in which successive values of S_i are correlated that generates the observed tent-shaped distribution. Suppose each firm has a tendency to maintain a value S^* , which evolves only slowly in time and which can be interpreted as the minimum point of a 'U-shaped' average cost curve in conventional economic theory. This type of dynamics is similar to what is known in economics as regression towards the mean^{27,28}. If the growth process has a constant 'back-drift', that is,

$$S_{i+\Delta t} / S_i = \begin{cases} k(1 + \varepsilon_i) & \text{for } S_i < S^* \\ \frac{1}{k}(1 + \varepsilon_i) & \text{for } S_i > S^* \end{cases} \quad (3)$$

where k is a constant larger than one, then the distribution of growth rates is the tent-shaped distribution equation (1) with a width proportional to $1/\ln k$ (ref. 29).

Our empirical findings of equation (2) are consistent with a hierarchical model of the internal structure of each firm. In zeroth-order approximation, suppose that a given company consists of independent units. If the unit's sales fluctuate with a standard deviation independent of s_0 , then equation (2) follows with $\beta = 1/2$. The much smaller empirical value of β that we find indicates the presence of strong, positive correlations among the firm's units. We propose a model relying on a technology of management (which may be common across firms) as opposed to a technology of production; this model may lead to some insight into why the behaviour of apparently diverse firms follows a simple law.

Consider a tree-like hierarchical organization of a firm¹⁷. The root of the tree (that is, 'top of the pyramid') represents the head of the firm, whose policy is processed to the level beneath, and so on, until finally the $N = z^n$ lowest-level units take action; here z is the average number of links connecting the levels and n the average number of levels. The N lowest-level units have sales ξ_i and mean $\langle \xi \rangle$, so $S_0 = \sum_{i=1}^N \xi_i = N \langle \xi \rangle$.

Suppose that the head of the company suggests a policy with the intention to change the sales of each lowest-level unit by an amount $\Delta \xi$. If this policy were to be propagated

through the hierarchy without any modifications, then the change in sales would be $\Delta S = N\Delta\xi = S_0\Delta\xi/\langle\xi\rangle$. Accordingly, $r = \ln[(S_0 + \Delta S)/S_0] = \ln[1 + \Delta\xi/\langle\xi\rangle]$, which is independent of S_0 . It follows that $\beta = 0$.

More realistically, each unit is not only influenced by the policy of the head but also by other (external and internal) factors. An example is that different levels have different information. Managers at each level might deviate from decisions made higher up in the tree if other information suggests to them that another action is appropriate. Another reason for a modification of the policy is organizational failure, due either to poor communication or disobedience. For these reasons, we assume that each manager follows his supervisor's policy with a probability Π , while the probability $(1 - \Pi)$ imposes a new independent policy for his subunits. Straightforward calculation using methods described, for example, in ref. 30 yields equation (2) for $n \gg 1$, with exponent β given by the formula

$$\beta = \begin{cases} -\ln \Pi / \ln z & \text{if } \Pi > z^{-1/2} \\ 1/2 & \text{if } \Pi \leq z^{-1/2} \end{cases} \quad (4)$$

(for small n , equation (4) is still a good approximation—for example, for $n = 3$ and $z = 2$, the deviation from the value $\beta = 0.20$ is only 0.03). Equation (4) is confirmed in the two limiting cases: when $\Pi = 1$ (absolute control) $\beta = 0$, while for all $\Pi < 1/z^{1/2}$, decisions at the upper levels of management have no statistical effect on decisions made at lower levels, and $\beta = 1/2$. Moreover, for a given value of $\beta < 1/2$ the control level Π will be a decreasing function of z : $\Pi = z^{-\beta}$. For example, if we choose the empirical value $\beta \approx 0.15$, then equation (4) predicts the plausible result $0.9 \geq \Pi \geq 0.7$ for a range of z in the interval $2 \leq z \leq 10$.

Our central results, equations (1) and (2), constitute a test that any accurate theory of the firm must pass. These equations support the possibility³¹ that the scaling laws used to describe complex systems comprised of many interacting inanimate particles (as in many physical systems) may be usefully extended to describe complex systems comprised of many interacting animate subsystems (as in economics). $\square$

Received 13 July 1995; accepted 31 January 1996.

- Gibrat, R. *Les Inégalités Économiques* (Sirey, Paris, 1933).
- Coase, R. H. *Economica* **4**, 386–405 (1937).
- Hart, P. E. & Prais, S. J. *J. R. statist. Ser. A* **119**, 150–191 (1956).
- Simon, H. A. & Bonini, C. P. *Am. Econ. Rev.* **48**, 607–617 (1958).
- Baumol, W. *Business Behavior, Value, and Growth* (MacMillan, New York, 1959).
- Hymer, S. & Pashigian, P. J. *polit. Econ.* **52**, 556–569 (1962).
- Cyert, R. & March, J. *A Behavioral Theory of the Firm* (Prentice-Hall, Englewood Cliffs, New Jersey, 1963).
- Jensen, M. C. & Meckling, W. H. *J. fin. Econ.* **3**, 305–360 (1976).
- Ijiri, Y. & Simon, H. A. *Skew Distributions and the Sizes of Business Firms* (North Holland, Amsterdam, 1977).
- Lucas, R. *Bell J. Econ.* **9**, 508–523 (1978).
- Jovanovic, B. *Econometrica* **50**, 649–670 (1982).
- Nelson, R. R. & Winter, S. G. *An Evolutionary Theory of Technical Change* (Harvard Univ. Press, Cambridge, Massachusetts, 1982).
- Golan, A. thesis, Univ. California, Berkeley (1988).
- Holmstrom, B. R. & Tirole, J. in *Handbook of Industrial Organization* Vol. 1 (eds Schmalensee, R. & Willig, R.) 61–133 (North Holland, Amsterdam, 1989).
- Williamson, O. E. in *Handbook of Industrial Organization* Vol. 1 (eds Schmalensee, R. & Willig, R.) 135–182 (North Holland, Amsterdam, 1989).
- Milgrom, P. & Roberts, J. *Economics, Organization, and Management* (Prentice-Hall, Englewood Cliffs, New Jersey, 1992).
- Radner, R. *Econometrica* **61**, 1109–1146 (1993).
- Golan, A. *Adv. Econometrics* **10**, 1–46 (1994).
- Ijiri, Y. & Hattiwanger, J. Q. *J. Econ.* **107**, 819–863 (1992).
- Davis, S. J., Hattiwanger, J. & Schuh, S. J. *Job Creation and Destruction* (MIT Press, Cambridge, Massachusetts, 1996).
- Singh, A. & Whittington, G. *Rev. Econ. Studies* **42**, 15–26 (1975).
- Hall, B. H. *J. Indust. Econ.* **35**, 583–606 (1987).
- Evans, D. S. *J. Polit. Econ.* **95**, 657–674 (1987).
- Shapiro, D., Bollman, R. D. & Ehrensaft, P. *Am. J. Ag. Econ.* **69**, 477–483 (1987).
- Dunne, T., Roberts, M. & Samuelson, L. Q. *J. Econ.* **104**, 671–698 (1989).
- Stanley, M. H. R. et al. *Econ. Lett.* **49**, 453–457 (1995).
- Friedman, M. *J. Econ. Lit.* **30**, 2129–2132 (1992).
- Leonard, J. Working Paper No. 1951 (National Bureau of Economic Research, Cambridge, MA, 1986).
- Risten, H. *The Fokker-Planck Equation* (Springer, Heidelberg, 1988).
- Li, W. *Phys. Rev.* **A43**, 5240–5260 (1991).
- Mantegna, R. N. & Stanley, H. E. *Nature* **376**, 46–49 (1995).

ACKNOWLEDGEMENTS. We thank R. N. Mantegna and J. Miron for help in the early stages of this work. This work was supported by JNICT (L.A.), DFG (H.L. and P.M.) and the US NSF.

Melting dynamics of a plasma crystal

Hubertus M. Thomas & Gregor E. Morfill

Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, 85740 Garching, Germany

PLASMAS have long been regarded as the most disordered state of matter; nevertheless, a set of colloidal particles introduced into a charge-neutral plasma can spontaneously exhibit ordered crystalline structures^{1,2}—so-called ‘plasma crystals’. Such systems, which reach equilibrium very rapidly and can be easily tuned between their ordered and disordered states, are ideally suited for investigating the processes underlying the solid-to-liquid phase transition. Here we report the results of experiments on ‘flat’ plasma crystals (with thicknesses of only a few lattice planes) which suggest that the melting transition occurs through two fundamental intermediate stages. On melting, the crystal first enters a state characterized by islands of crystalline order, about which streams of particles flow. The crystalline regions then dissolve as the vibrational energy of the system increases, but this is accompanied by a temporary increase in orientational order before the system finally enters a disordered, liquid state. The unexpected ‘vibrational’ phase, characterized by enhanced orientational order, might arise as a consequence of the mixed two- and three-dimensional nature of the flat plasma crystals. Alternatively, it may indicate the existence of a new intermediate state in melting transitions more generally.

The forces controlling the structure and thermodynamics of plasma crystals are (1) Coulomb forces between the embedded particles and (2) neutral gas friction, which ‘cools’ the particles down to brownian motion. We consider Coulomb forces first. By colliding with electrons and ions in a plasma, micrometre-sized particles may obtain (negative) equilibrium charges, Q , of several thousand elementary charges, e (refs 1, 3). Equilibrium is reached in a fraction of a second. The plasma reorganizes itself locally in the electric field around the particle to neutralize its charge. The appropriate ‘screening distance’ is called the Debye length, λ_D ; it depends on the plasma temperature, T , and the plasma density, n , as $\lambda_D = (kT/4\pi n e^2)^{1/2}$. Coulomb interaction between neighbouring particles (and therefore crystallization) can only occur if their separation is $\lesssim \lambda_D$. (This system may be considered to be comparable to a heavy atom whose positive core is neutralized by its electron cloud. The effective ‘atom radius’ is λ_D .) The second controlling force is neutral gas friction, which can provide the low particle kinetic energies required for crystallization. Experimental conditions for crystallization thus require a partially ionized plasma with a low- (room-) temperature neutral gas component. Such conditions can be produced easily in low-power radio-frequency discharges. For the experiment described here, we used krypton at room temperature and pressures between 0.1 and 0.5 mbar, a radio-frequency power of 0.8–2 W to give an ionization fraction from 10^{-7} to 10^{-6} , and monodisperse melamine/formaldehyde spheres of 6.9 μm diameter (see ref. 1 for details).

The parameters determining the thermodynamics of the plasma crystal system are (1) the Coulomb coupling parameter Γ . In the case of a screened Debye potential, $\Gamma \equiv (Q^2/\Delta kT) \exp(-\Delta/\lambda_D)$. This is the ratio of the Coulomb energy between two neighbouring particles (separated by a distance Δ) to their kinetic energy kT . (2) The ratio $\kappa = \Delta/\lambda_D$ of the lattice distance, Δ , to the Debye screening distance. Theoretical considerations^{4,5} suggest that Coulomb crystallization may occur if $\Gamma \geq \Gamma_c = 172$ and $\kappa \lesssim 1$. (Γ_c is the critical value for crystallization obtained in Monte Carlo simulations of one-component-plasmas.) A decrease of Γ , or an increase of κ , leads to ‘melting’ of the crystalline structure. This can be initiated and controlled easily in radio-frequency discharge

plasmas by changing the gas pressure.

There are other model crystal systems (apart from the natural ones) with which phase transitions may in principle be studied. Wigner⁶ pioneered the field with the theory of the so-called "Wigner crystal". Since then, technological advances in low-temperature physics and quantum optics have led to one-component-crystals such as ion crystals^{7,8,9} and electron crystals^{10,11}. In addition there are the colloidal suspensions in aqueous solutions¹²⁻¹⁴, which have been most widely used for phase-transition studies. For a recent overview see ref. 15.

Comparing plasma crystals with the above systems, we note the similarities (and some major differences) with the colloidal suspensions. Both are overall charge-neutral systems and utilize small (colloidal) particles embedded in a partially ionized medium to organize and visualize structures produced by Coulomb interactions. The unique properties of plasma crystals are: very fast response times (more than a million times faster than colloidal suspensions), very little damping (about six orders of magnitude smaller than colloidal suspensions), easy experimental control, detailed imaging and high time resolution of the dynamics of individual particles, both in two and three dimensions. These properties make plasma crystals particularly useful for studying

dynamical processes (for example, the solid/liquid phase transition, self-organization, and coherent and incoherent perturbations).

We have produced monolayers as well as three-dimensional plasma crystals, with more than a hundred lattice distances in the horizontal plane and up to 18 layers in the vertical direction. Gravity acts to restrict the vertical lattice planes to a few¹⁶. Bigger three-dimensional crystals require either microgravity conditions¹⁷ or, possibly, smaller particles¹⁶.

In our experiments, plasma crystal structures are visualized by illuminating each lattice layer with a sheet of laser light and viewing with a CCD (charge-coupled device) camera normal to this sheet. Particle positions are stored on a video recorder with a high temporal and spatial resolution. Three-dimensional structures are determined by vertical stepping. The 'melting experiment' starts from a well established crystalline state at 0.42 mbar pressure; transitions are followed by continuous lowering of the gas pressure. Both structural and dynamical properties of the plasma crystal are determined, for example, translational and bond-orientational correlation functions, random (thermal) and systematic particle motion.

From these analyses we identify the following 'states' during the melting transition.

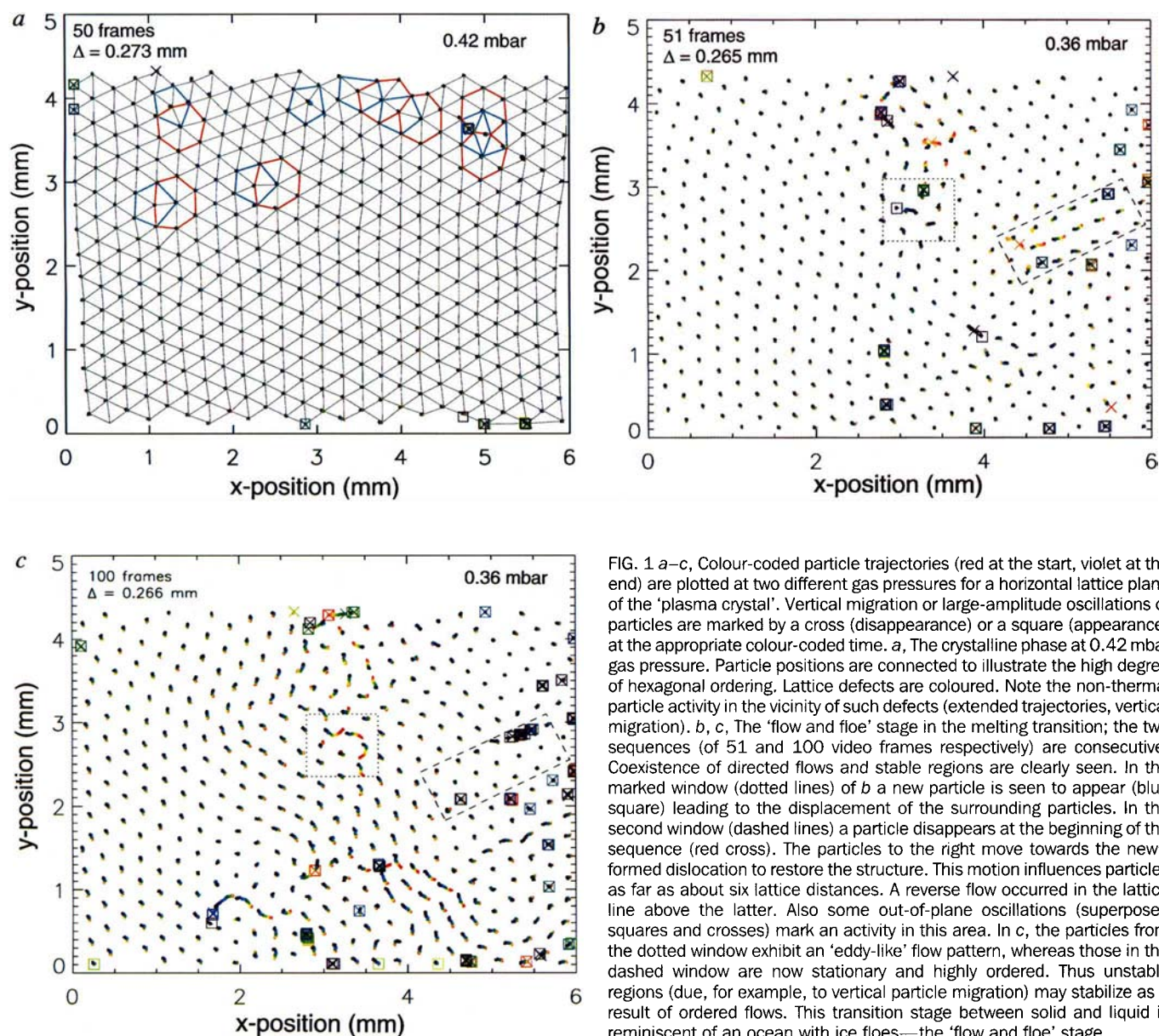


FIG. 1 *a-c*, Colour-coded particle trajectories (red at the start, violet at the end) are plotted at two different gas pressures for a horizontal lattice plane of the 'plasma crystal'. Vertical migration or large-amplitude oscillations of particles are marked by a cross (disappearance) or a square (appearance) at the appropriate colour-coded time. *a*, The crystalline phase at 0.42 mbar gas pressure. Particle positions are connected to illustrate the high degree of hexagonal ordering. Lattice defects are coloured. Note the non-thermal particle activity in the vicinity of such defects (extended trajectories, vertical migration). *b, c*, The 'flow and floe' stage in the melting transition; the two sequences (of 51 and 100 video frames respectively) are consecutive. Coexistence of directed flows and stable regions are clearly seen. In the marked window (dotted lines) of *b* a new particle is seen to appear (blue square) leading to the displacement of the surrounding particles. In the second window (dashed lines) a particle disappears at the beginning of the sequence (red cross). The particles to the right move towards the newly formed dislocation to restore the structure. This motion influences particles as far as about six lattice distances. A reverse flow occurred in the lattice line above the latter. Also some out-of-plane oscillations (superposed squares and crosses) mark an activity in this area. In *c*, the particles from the dotted window exhibit an 'eddy-like' flow pattern, whereas those in the dashed window are now stationary and highly ordered. Thus unstable regions (due, for example, to vertical particle migration) may stabilize as a result of ordered flows. This transition stage between solid and liquid is reminiscent of an ocean with ice floes—the 'flow and floe' stage.

'Crystalline', characterized by high translational and orientational order, few lattice defects, energy of lattice vibrations $\sim kT$ —with T at approximately room temperature, occasional (rare) large-amplitude nonthermal oscillations—mostly in the immediate vicinity of lattice defects.

'Flow and floe', characterized by the coexistence of islands of ordered crystalline structure (floes) and systematic directed particle motion (flows), thermal motion corresponds to room temperature ($v_{th} \approx 0.2 \text{ mm s}^{-1}$)—directed flow velocities are typically half of this—translational and orientational order have decreased significantly, occasional vertical particle migration to other lattice planes.

'Vibrational', characterized by a return to a more orientationally ordered structure and diminishing flow-regions, increasing vibrational amplitudes, thermal energy and vertical migration. The translational order continues to decrease.

'Disordered', characterized by collisions, complete vertical as well as horizontal migration, no discernible translational or orientational order, thermal energy $\geq 200 \times$ room temperature, T of the order unity.

We now illustrate the different transition 'states' described above with the data from our experiment. Figure 1a shows the crystalline phase. The observed particle positions have been connected to emphasize the hexagonal crystal structure. Defects are marked and particles with non-thermal behaviour are also identified. Figure 1b and c shows examples of the 'flow and floe' state. Coexistence of directed systematic flows and islands of stable crystal structure are clearly seen. Such behaviour, which is well-known in three-dimensional systems, of course, has been recognised in some two-dimensional colloidal systems as well^{18,19}, where it was noted that ordered flows are associated with lattice defects and rearrangement of dislocations. The analysis of plasma crystals sheds new light on the physical origin. We note from Fig. 1b and c (as well as many other examples not shown) that ordered flows may be triggered by vertical particle migration. Because the

'flat' crystal investigated here has a hexagonal structure in its horizontal planes and is 'stacked' vertically (such that particles are aligned), it appears that energetically it is easier to compensate vertical migration by horizontal shear flows—creating conditions for returning a 'lost' particle at an energetically more favourable location elsewhere. In this way the crystal exhibits complicated 'convection patterns', possibly connected over many lattice planes, which characterize this 'flow and floe' stage.

Figure 2 shows portions of the original video images. Figure 2a and b corresponds to the middle and end of the 'vibrational state'. The flow phenomenon, so prominent in Fig. 1b and c, is not very important. The major characteristics of this state are, rather surprisingly, a growth in the orientational order and an increase in the thermal (vibrational) energy (see Figs 3 and 4). Migration across lattice planes is more frequent, presumably due to the increase in thermal energy. This could also be responsible for the suppression of ordered flows—greater vertical mobility obviates the need for adjustments through horizontal shear flows. The thermal energy of the particles grows as the gas pressure is reduced further. Figure 2c shows collisions, resulting in abrupt velocity changes and complete three-dimensional thermalization. Superficially, this state—called 'disordered'—might qualify as a gaseous state, although this has not been proved rigorously.

Figure 3 shows translational and orientational order parameters as a function of gas pressure. The translational order continuously decreases with decreasing gas pressure, whereas the orientational order develops a secondary maximum, which coincides with the state we have called 'vibrational'. In this state, the particle surface density remains constant, within the measurement uncertainties, at a somewhat higher (factor 1.06) value than the crystalline state. Thus the system has an 'equation of state' where the kinetic energy of the particles may increase by a factor ~ 100 at constant density!

In Fig. 4 the particle velocity distributions for the four states, namely crystalline, flow and floe, vibrational and disordered are

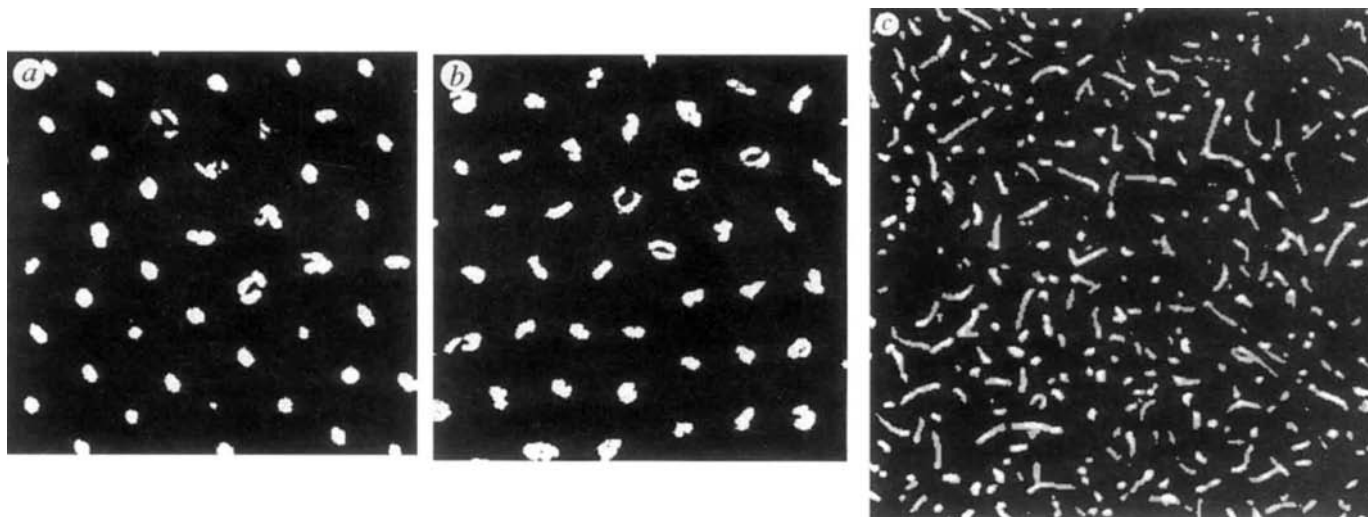


FIG. 2 a–c, Examples of video images, demonstrating some of the features of the 'vibrational' and 'disordered' states of the melting transition. (The name 'vibrational state' was chosen, because it highlights the major aspect of the observed particle dynamics during this transition phase.) a, Superposition of three video images (taken 0.02 s apart), showing the trajectories of individual particles at high spatial resolution. The area is $1.57 \times 1.63 \text{ mm}$ and the pressure 0.32 mbar. Note the circular trajectories, as well as the elongated ellipsoidal images and the dots. These could all be different projections of the particle oscillations. Note also the high degree of orientational (six-fold) order. b, Taken at 0.29 mbar (at the end of the vibrational stage) also shows the large-amplitude oscillations, that characterize this state, as well as the orientational order (although this is

weaker than a). The dimension and number of superposed video images of this panel are the same as those of a, so that the two can be directly compared. The oscillation amplitudes have grown to significant fractions of the lattice separation, leading to more frequent migrations across lattice planes. The large-scale view of this crystal plane (not shown) indicates that the regions with high vibrational energy start in a few restricted locations and then spread throughout the crystal as the gas pressure is lowered. c, Taken at 0.13 mbar, is just a single video image (exposure time 0.02 s). Note the long trajectories, the sudden directional changes in particle motion resulting from particle–particle collisions and the high level of disorder. The translational (pair correlation) order parameter at this stage shows no structure at all—a necessary, but not sufficient, requirement for the gas phase.

FIG. 3 The evolution of translational and orientational order parameters during the observed phase transition, as a function of neutral gas pressure. The radial pair distribution function $g(r)$ is fitted for an ideal hexagonal (or triangular) lattice with a lattice constant, Δ , (the mean obtained from the data), broadened by a Debye–Waller factor with scale b , and an exponential decay with a decay constant η —the translational order parameter shown here (left-hand y-axis) is the dimensionless quantity $\eta\Delta$ (triangles). For details see, for example, ref. 13. High translational order corresponds to small decay constants. The bond orientational correlation function is $g_6(r)$, which is defined in terms of the nearest-neighbour bond angles and measures the six-fold symmetry in the structure²⁰. Shown here (right-hand y-axis) is $g_6(0)$ averaged over all the particles, that is, the mean local orientational order. High orientational order corresponds to large $g_6(0)$ —the maximum (ideal) value is 1. Both these order parameters can be used to identify the observed instantaneous structure with the prediction of the KTHNY two-dimensional melting theory^{21–25}. We have shaded three regions: at ~ 0.42 mbar, the crystalline phase; at ~ 0.32 mbar, the ‘vibrational’ state; and below ~ 0.23 mbar the ‘disordered’ state. The ‘flow and floe’ state occurs at ~ 0.36 mbar in this particular experimental configuration. Particle positions are calculated from the geometric mean of the distribution of bright pixels. The accuracy of this technique gets progressively worse, the higher the particle velocities.

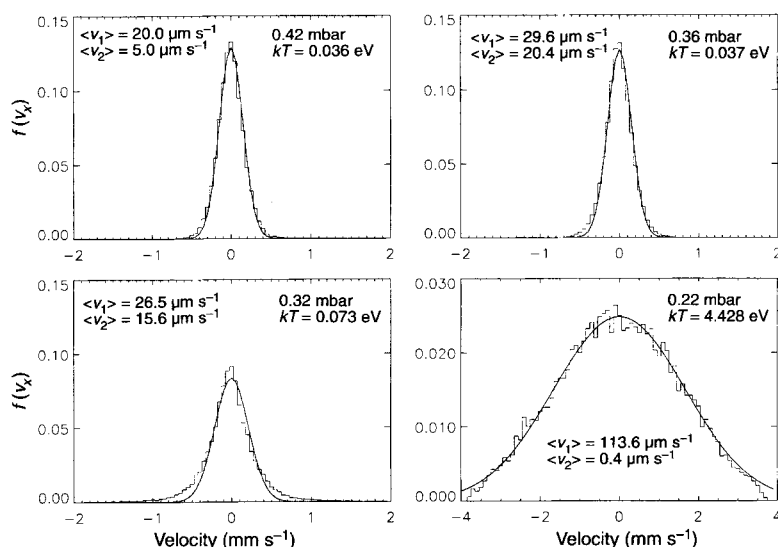
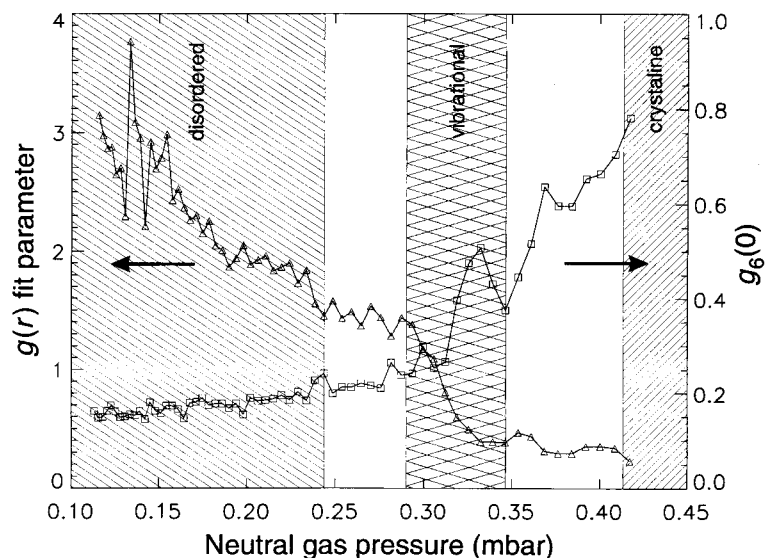


FIG. 4 Histograms of the distribution of particle velocity components in the x-direction, v_x , at four representative pressures (corresponding to the crystalline, flow and floe, vibrational and disordered states). The continuous curves represent the best-fit Maxwellian distribution, the corresponding thermal energy, kT , is indicated. Similar fits were made to the y-component (v_y) and the total particle velocity distribution (not shown). There was only a minor asymmetry in the velocity distributions. To quantify the occurrence of directed flows, which are so apparent in Fig. 1b and c, we have calculated two ‘drift’ velocities $\langle v_1 \rangle \equiv \langle |\Delta x_n/n\Delta t| \rangle$ and $\langle v_2 \rangle \equiv \langle \delta(y) |\Delta x_n/n\Delta t| \rangle$. Here Δx_n is the length of a particle’s trajectory $(\Delta x^2 + \Delta y^2)^{1/2}$ over n time steps (video frames) of duration Δt , and $\delta(y)$ is the δ -function ($= +1$ if the trajectory points in the positive y-direction, -1 if it points in the negative y-direction). Thus $\langle v_1 \rangle$ gives a general measure of ‘undirected’ motion, whereas $\langle v_2 \rangle$ gives a measure of net flow activity in the y-direction (or any other direction chosen). If $\langle v_1 \rangle$ is due to random motion, $\langle v_2 \rangle \ll \langle v_1 \rangle$, if there is a dominance of directed flow with a net y-component, $\langle v_2 \rangle < \langle v_1 \rangle$. We see from the examples shown here that the ‘flow and floe’ state at ~ 0.36 mbar satisfies the latter inequality better than the other states.

shown. The ‘vibrational state’ would be fitted better with a bi-Maxwellian distribution—some particles still have low thermal energies, some are thermally agitated.

We have described the solid-to-liquid (and possibly gaseous) phase transition for a plasma crystal. The transition appears to

develop via the growth of crystal defects, directed particle flows around crystalline islands (floes), enhanced vibrations associated with greater order and flow suppression, to a completely disordered state. This seems neither a straightforward first-order nor a classical second-order transition

Received 1 August 1995; accepted 18 January 1996.

1. Thomas, H. et al. *Phys. Rev. Lett.* **73**, 652–655 (1994).
2. Chu, J. & Lin, I. *Phys. Rev. Lett.* **72**, 4009–4012 (1994).
3. Melzer, A., Homann, A. & Piel, A. *Phys. Rev. E* (submitted).
4. Ikezi, H. *Phys. Fluids* **29**, 1764–1766 (1986).
5. Farouki, R. & Hamaguchi, S. *Appl. Phys. Lett.* **61**, 2973–2975 (1992).
6. Wigner, E. *Phys. Rev.* **46**, 1002–1011 (1934).
7. Diedrich, F., Peik, E., Chen, J., Quint, W. & Walther, H. *Phys. Rev. Lett.* **59**, 2931–2934 (1987).
8. Gilbert, S., Bollinger, J. & Wineland, D. *Phys. Rev. Lett.* **60**, 2022–2025 (1988).
9. Waki, I., Kassner, S., Birk, G. & Walther, H. *Phys. Rev. Lett.* **68**, 2007–2010 (1992).
10. Jiang, H. et al. *Phys. Rev. Lett.* **65**, 633–636 (1990).
11. Goldman, V., Santos, M., Shayegan, M. & Cunningham, J. *Phys. Rev. Lett.* **65**, 2189–2192 (1990).
12. Kose, A., Ozaki, M., Takano, K. & Kobayashi, Y. & Hachisu, S. *J. Colloid Interf. Sci.* **44**, 330–338 (1973).

13. Murray, C. *Bond Orientational Order in Condensed Matter Systems* (ed. Strandburg, K.) Ch. 4 (Springer, New York, 1992).
14. Grier, D. & Murray, C. *J. chem. Phys.* **100**, 9088–9095 (1994).
15. Maddox, J. *Nature* **370**, 411 (1994).
16. Morfill, G. E. & Thomas, H. M. *J. Vac. Sci. Technol. A* (in the press).
17. Morfill, G. E. et al. *Colombus Precursor Flights Proposal* (MPE preprint, 1991).
18. Murray, C. & Wenk, R. *Phys. Rev. Lett.* **58**, 1643–1646 (1989).
19. Hurley, M. & Harrowell, P. *Phys. Rev. E* **52**, 1694–1698 (1995).
20. Löwen, H. *J. Phys.: Condens. Matter* **4**, 10105–10116 (1992).
21. Kosterlitz, J. & Thouless, D. *J. Phys.* **C6**, 1181–1203 (1973).
22. Halperin, B. & Nelson, D. *Phys. Rev. Lett.* **41**, 121–124 (1978).
23. Nelson, D. & Halperin, B. *Phys. Rev.* **B19**, 2457–2484 (1979).
24. Young, A. *Phys. Rev.* **B19**, 1855–1857 (1979).
25. Nelson, D. in *Phase Transitions and Critical Phenomena* Vol. 7 (eds Domb, C. & Leibowitz, J.) (Academic, London, 1983).

ACKNOWLEDGEMENT. We thank L. Ratke for helpful discussions.

Contrasting atmospheric and climate dynamics of the last-glacial and Holocene periods

Peter D. Ditlevsen*, Henrik Svensmark† & Sigfus Johnsen*‡

*The Niels Bohr Institute, Department for Geophysics, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark

†Solar-Terrestrial Physics Division, Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen Ø, Denmark

‡Science Institute, Department of Geophysics, University of Iceland, Dunhaga 3, IS-107, Reykjavik, Iceland

Our present climate is relatively stable compared to that of the Last Glacial Maximum about 20,000 years ago. Palaeoclimate records obtained from ice cores^{1,2} and deep-sea sediment cores³ for the last glacial period show abrupt temperature changes on timescales of a few hundred years, which have been attributed to cycles of ice build-up and release associated with large ice sheets (Dansgaard-Oeschger cycles and Heinrich events)³ and their coupling to ocean circulation^{4,5}. But little is known about the dynamics of the atmosphere during the last glaciation. Ice sheets influence atmospheric circulation, and studies using general circulation models have suggested stormier, more variable atmospheric dynamics during the Last Glacial Maximum than today⁶⁻⁹. Here we report the results of an analysis of temporal trends over the past 91,000 years in the oxygen isotope signatures of a high-resolution ice-core record from Greenland^{1,2}. This analysis provides direct evidence that atmospheric circulation during the last glaciation was more turbulent than it is today.

Climate dynamics are characterized by a variety of physical processes covering a large range of timescales. To retrieve information on the important physical processes it is crucial to obtain and analyse climate signals covering these timescales. The $\delta^{18}\text{O}$ isotope data extracted from the ice core provide, owing to the high snow accumulation rate in central Greenland, a high-resolution climate signal as a proxy for the temperature (see Fig. 1 legend).

The ice-core data analysed here covers the Holocene and last-glacial periods with a temporal resolution ranging from seasons in the present to ~10 yr at 91 kyr before present, BP (17,496 data points). The data is based on sampling slices of specified increments down through the ice core. The connection between the depth in the core and the age is established by counting of annual layers down to 14 kyr BP and a model for the ice-flow dynamics further down¹⁰. The full signal is shown in Fig. 1a. The power density spectrum is shown in Fig. 2.

There are two regimes of behaviour in the power spectrum separated at around a few hundred years. For timescales longer than 100–200 yr the spectrum is continuous, without dominant peaks, of the form $1/f^\alpha$ with $\alpha \approx 1.6$. At this point the spectral slope (~ -1.6) of the low-frequency part of the signal diminishes to ~ 1 (white noise). To separate the climate information of these two regimes we split the signal into a high- and a low-frequency part using a spectral cut at 150 years; Fig. 1b shows the 150-yr low-pass, and Fig. 1c is the 150-yr high-pass. To this we have applied a 20-yr low-pass to remove artificial trends in the record caused by a poorer resolution in the bottom of the record.

The characteristic “saw-tooth” variations of the low-pass filtered signal, of typical timescales around a few thousand years, have previously been assigned to the ice/ocean dynamics through ice-surges (Dansgaard-Oeschger and Heinrich events)³. The climate dynamics of these events are probably related to surging and build-up of the ice-caps, and coupling to the ocean circulation^{4,5}. The surging is a very rapid process whereas the build-up takes more than a few hundred years, as the ice has to be

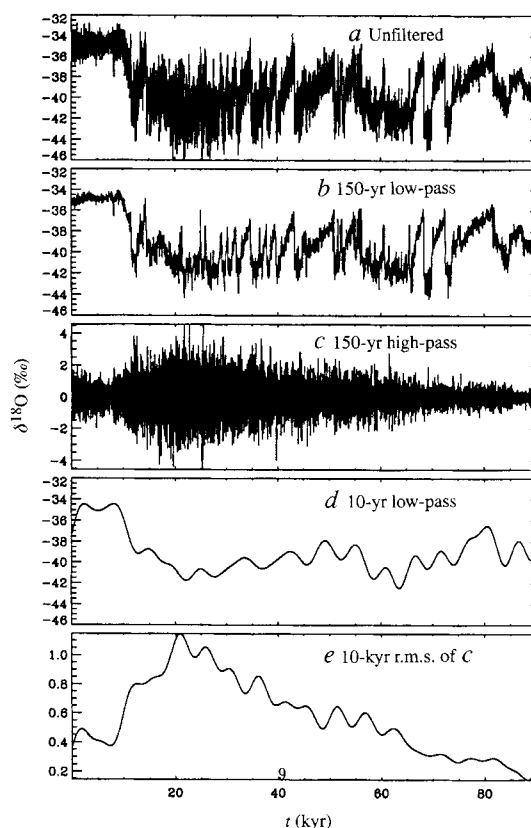


FIG. 1 a, The high-resolution $\delta^{18}\text{O}$ signal from the Greenland ice-cap. $\delta^{18}\text{O}$ is given (in units of ‰) by $[(^{18}\text{O}/^{16}\text{O})_{\text{ice}}/(^{18}\text{O}/^{16}\text{O})_{\text{standard}}] - 1$, where the standard is present-day ocean water. Empirically it has been shown that there is approximately a constant linear relationship between $\delta^{18}\text{O}$ in precipitation and the air temperature where the precipitation occurs¹⁹. The ice-cap thus constitutes a record which is a proxy for air temperature. The exact relationship between temperature and isotope ratio depends in an intricate way on the sea/air temperature of evaporation, cloud temperature of condensation, and the path followed by the vapour before falling out as snowflakes. Furthermore, the isotope ratio in the ice depends on molecular diffusion and mixing in the ice. The ice-core $\delta^{18}\text{O}$ ratio is therefore a proxy for some spatially and temporally averaged temperature signal, conditioned by a precipitation event. But we expect that these processes act as independent noise on the climate temperature signal contained in the record. b, The 150-yr low-pass of the $\delta^{18}\text{O}$ signal; c, the corresponding 150-yr high-pass. d, The 10-kyr low-pass of the $\delta^{18}\text{O}$ signal; and e, the 10-kyr low-pass of the absolute value of the 150-yr high-pass (c), corresponding to the 10-kyr root mean square (r.m.s.) of the high-pass.

transported as precipitation through the atmosphere. This could explain the saw-tooth shapes, which are maintained in the low-pass signal, Fig. 1b. There is at present no full understanding of the connection between the complex dynamics underlying the signal and the spectral characteristics. But the slope (-1.6) of the power spectrum indicates non-trivial dynamics over long timescales, in the sense that the signal cannot be explained as a simple autoregressive process. In a simple autoregressive process, as that of a stochastic climate model¹¹, the signal is generated by accumulation of independent identically distributed noise, representing the fast components of the climate system. The power spectrum will, for such a process, on timescales shorter than the correlation time always have a spectral slope of -2 .

The residual high-pass signal (Fig. 1c) represents timescales faster than a few hundred years. This part of the signal contains the information on the atmospheric dynamics. The most striking feature of this signal is that the envelope of the fluctuations is roughly proportional to the degree of glaciation—or the tempera-

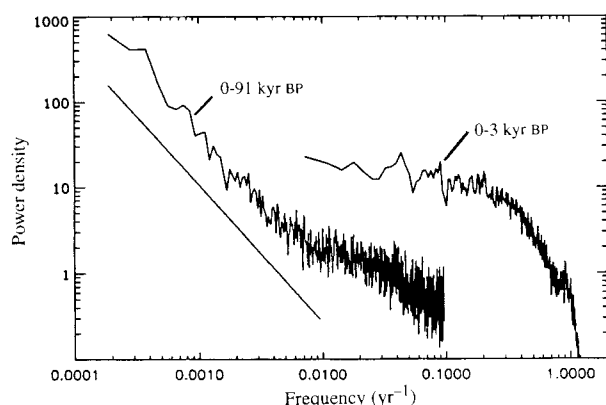


FIG. 2 Power density spectra as a function of frequency for the $\delta^{18}\text{O}$ record covering 0–91 kyr BP with a temporal resolution of 10 yr at 91 kyr BP (17,496 points). The rightmost, Holocene, spectrum covers 0–3 kyr BP, with a temporal resolution of approximately one month (26,244 points). The spectra have been calculated from the temporal signals using a sliding window. The spectrum has a slope of ~ -1.6 for timescales larger than 100–200 yr, indicated by the line. For timescales smaller than 100–200 yr the spectrum is white until timescales of 5–10 yr. For the Holocene spectrum the annual peak is observed. We see no signs of significant peaks, except for the annual, above the noise level of continuous background in the spectrum. We have used sliding windows of different sizes to look for peaks, but we do not expect an oscillatory behaviour of the climate system on the timescales smaller than those of the orbital forcing. For the orbital periods the record is not long enough for a proper spectral analysis. There are other techniques like singular spectrum analysis (SSA) and the multi-taper method (MTM) designed to extract small peaks from the continuous background, but one must be very careful when applying these techniques, as they tend to separate realizations of processes with a continuous power spectrum into peaks where the significance tests are rather doubtful, and depend on which null hypothesis is assumed.

ture as represented by the $\delta^{18}\text{O}$ itself. Figure 1d shows the 10-kyr low-pass of $\delta^{18}\text{O}$ (Fig. 1a), and Fig. 1e shows the 10-kyr low-pass of the absolute value of the high-pass (Fig. 1c). The $\delta^{18}\text{O}$ signal is a proxy for the local temperature; the relation between the two quantities is that in a cold climate, storm tracks move southwards and the transport route for the precipitation is longer increasing the depletion of ^{18}O . Thus we interpret the increased variance as a direct result of a more stormy—or turbulent—state of the atmospheric flow during the Last Glacial Maximum (LGM).

As an independent measure of the atmospheric and climatic variability we have analysed the probability density functions (PDFs) of the high-pass signal which, in contrast to the low-pass signal, is stationary. The power spectrum of a signal contains information on the second moments and the correlation structure of the signal, whereas the PDFs contains information on all moments but no information on the temporal correlation structure.

A characteristic of the state of turbulent atmospheric flow is the intermittency or occurrence of extreme values in the temperature field as represented through the PDFs¹². In present-day temperature records, gaussian PDFs are found when subtracting the annual and diurnal cycles. Accordingly, we have analysed the ice-core record for the past 3 kyr, with a temporal resolution of approximately one month (26,244 points), referred to as the Holocene signal in the following. This is compared with the signal covering 14.4–29 kyr BP with a temporal resolution of ~ 3.7 yr (3,888 points), referred to as the LGM signal in the following. The Holocene signal has a white-noise power spectrum down to 5–10 yr (Fig. 2) where it bends off. This is in agreement with spectra obtained from meteorological data^{13,14}. Note the annual cycle in the spectrum. In order to subtract the low-frequency variability, related to non-atmospheric components of the climate system, we

examine the high-pass filtered signal, and examine the PDF as a function of the high-pass cutoff frequency. The high-pass signal is normalized with the running variance in a window of 500 data points. The PDF are fitted to Laplace-type distributions¹⁵ of the form $\phi(x) = 1/(2^{1-1/\beta}\Gamma(1+1/\beta)\sigma) \exp(-|x/\sigma|^\beta/2)$, which for $\beta = 2$ is the normal distribution and for $\beta = 1$ is the Laplace distribution, Γ is the gamma-function and σ is related to the variance by $\text{var}(x) = \sigma^2 2^{2/\beta}\Gamma(3/\beta)/\Gamma(1/\beta)$. Figure 3a shows the normalized 30-yr high-pass for the Holocene signal, Fig. 3b shows the cumulated distribution on a normal probability scale, and Fig. 3c shows the fitted PDF. This signal is very close to being gaussian. Figure 3d, e and f shows the same for the normalized 30-yr high-pass of the LGM signal. In this case there is a significant deviation from a gaussian distribution. The kurtosis, the normalized fourth-order moment—a measure of the distributions deviation from being gaussian—and the best-fit $\beta = \beta(\kappa)$, signifying the Laplace-type distribution parameter, are shown in Fig. 4 as a function of the high-pass cutoff frequency. Triangles represent Holocene data, and diamonds LGM data; the error bars represent the 95% confidence level¹⁶. Again the present Holocene climate shows a gaussian distribution (kurtosis = 3, $\beta = 2$) for all time-scales, whereas the LGM signal becomes more intermittent for the faster time-scales.

Owing to the presence of the ice-sheets and a substantially colder northern ocean¹⁷ in the glacial period, larger thermal gradients between Equator and the glaciers are expected than in the present climate. This would cause a more energetic and turbulent atmosphere in the glacial climate. In the glacial period the distribution of the flat part of the power density spectrum, corresponding to timescales shorter than 100–200 yr, is not trivially gaussian. We believe that the increased variance and—independently—the observed PDFs in the glacial period are observations

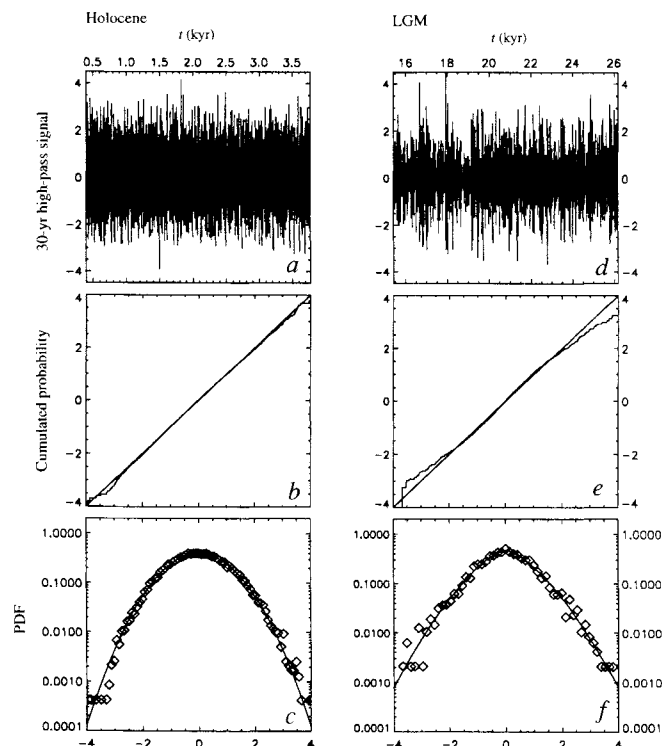


FIG. 3 a, The 30-yr high-pass filtered Holocene signal, covering 0–3 kyr BP (26,244 points). The signal is normalized with the variance in a 500-point running window; b the cumulated probability on a normal probability scale. The straight line corresponds to the gaussian distribution; c, the PDF, which is gaussian. d–f, As a–c, but for the LGM signal, covering 14.4–29 kyr BP (3,888 points). The signal is intermittent.

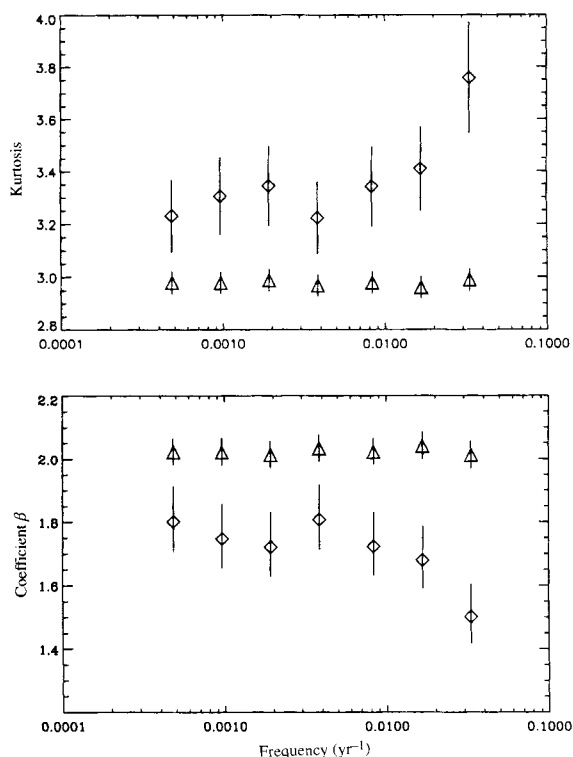


FIG. 4 Top, the kurtosis, $\kappa \equiv E(x^4)/E(x^2)^2$, as a function of the high-pass cutoff frequency. The kurtosis is 3 for the gaussian distribution and 6 for the laplacian distribution. Bottom, the best-fit parameter β (see text) as a function of the high-pass cutoff. For $\beta = 2$ the distribution is gaussian; for $\beta = 1$ it is laplacian. Triangles, Holocene; diamonds, LGM. The error bars are the 95% confidence levels. The Holocene signal is gaussian distributed for all cutoffs, whereas the LGM signal becomes more intermittent when subtracting the longer timescales.

of this more turbulent state of the atmospheric flow reflected in the $\delta^{18}\text{O}$ signal in Greenland. This is a direct observation of changes in the atmospheric circulation from palaeoclimate records for the LGM. It supports the indications of circulation changes in glacial climate inferred from correlations between snow accumulation and $\delta^{18}\text{O}$ correlations (ref. 18). □

Received 3 April 1995; accepted 24 January 1996.

1. Dansgaard, W. et al. *Nature* **264**, 218–220 (1993).
2. GRIP members *Nature* **364**, 203–207 (1993).
3. Bond, G. et al. *Nature* **365**, 143–147 (1993).
4. Bond, G. C. & Lotti, R. *Science* **267**, 1005–1010 (1995).
5. Fronval, T., Jansen, E., Bloemendal, J. & Johnsen, S. *Nature* **374**, 443–446 (1995).
6. Manabe, S. & Broccoli, A. J. *J. Geophys. Res.* **90**, 2167–2190 (1985).
7. Kutzbach, J. E. & Wright, H. E. *Quat. Sci. Rev.* **4**, 147–187 (1985).
8. Rind, D. J. *Geophys. Res.* **92**, 4241–4281 (1987).
9. Valdes, P. & Hall, (eds Duplessy, J.-C. & Spyridakis, M.-T.) (NATO ASI ser. 22, Springer, Berlin, 1994).
10. Dahl-Jensen, D. et al. (ed. Peltier, W. R.) 517–532 (NATO ASI ser. I 12, Springer, Berlin, 1993).
11. Hasselmann, K. *Tellus* **28**, 473–485 (1976).
12. Wu, X. & Libchaber, A. *Phys. Rev.* **A40**, 6421–6430 (1989).
13. Willebrand, J. J. *phys. Oceanogr.* **8**, 1080–1094 (1978).
14. Chave, A. D., Luther, D. S. & Filloux, J. H. *J. geophys. Res.* **96**, 18361–18379 (1991).
15. Johnson, N. L. & Kotz, S. *Distributions in Statistics: Continuous Univariate distributions II* (Wiley, New York, 1970).
16. Kendall, M. et al. *Kendall's Advanced Theory of Statistics* (Griffin, London, 1987).
17. Martinson, D. G. et al. *Quat. Res.* **27**, 1–29 (1987).
18. Kapsner, W. R., Alley, R. B., Shuman, C. A., Anandakrishnan, S. & Grootes, P. M. *Nature* **373**, 52–54 (1995).
19. Johnsen, S. J., Dansgaard, W. & White, J. W. C. *Tellus* **B41**, 452–468 (1992).

ACKNOWLEDGEMENTS. We thank J.-P. Steffensen, O. Ditlevsen and P. Dimon for discussions. The Greenland Ice-core Project (GRIP) is a European Science Foundation programme with eight nations collaborating in ice-core drilling. P.D.D. thanks the Carlsberg Foundation for funding.

Correlated progression and the origin of turtles

Michael S. Y. Lee

Zoology Building A08, School of Biological Sciences, University of Sydney, New South Wales 2006, Australia

TURTLES exhibit some of the most extreme postcranial modifications found in vertebrates. The dorsal vertebrae and ribs have fused with dermal armour, forming a totally rigid box-like trunk region^{1,2}. Our understanding of chelonian origins has been restricted by a paucity of information on intermediate forms^{3,4}, however, and it is often assumed that they must have evolved saltatorially⁵. It has been suggested that pareiasaurs, a group of large herbivorous anapsid reptiles, are the sister-group of turtles⁶. Here I show that certain pareiasaurs—dwarf, heavily armoured forms such as *Nanoparia*—approach the chelonian morphology even more closely than previously thought. Evolutionary trends within pareiasaurs, such as the elaboration of the dermal armour, shortening and stiffening of the presacral region, and increased reliance on limb-driven as opposed to axial-driven locomotion, suggest that the rigid armoured body of turtles evolved gradually, through 'correlated progression'⁷.

A consensus has recently been reached that turtles are part of a radiation of primitive reptiles known as 'parareptiles'^{8,9}, although there is still debate over which 'parareptiles' are closest to turtles: procolophonoids⁸ or pareiasaurs⁹. A comprehensive phylogenetic analysis of basal amniotes was therefore undertaken in order to resolve the external relationships of turtles¹⁰. Ingroup taxa included nycteroleterids, nyctiphuretid, procolophonids, *Barasaurus*, *Owenetta*, *lanthanosuchids*, *Sclerosaurus*, 'pareiasaurs' and turtles. Because evidence for the monophyly of the Pareiasauridae was weak, each species of pareiasaur was treated as a separate terminal taxon. Previous studies^{8,9} had treated pareiasaurs as a single discrete terminal taxon. All valid characters employed in previous analyses were included, although many new informative characters (mostly postcranial) were identified. The two nearest outgroups (millerettids and 'eureptiles'⁸) were used to polarize characters. The data were analysed using the branch and bound algorithm of PAUP 3.1.1.

The results confirm that pareiasaurs are the nearest relatives of turtles, although the exact nature of this relationship remains equivocal. In the most parsimonious tree (221 steps, c.i. = 0.77, r.i. = 0.92), pareiasaurs are ancestral to (that is, paraphyletic with respect to) turtles (Fig. 1). Large (2.5 m long, excluding tail) heavily ossified pareiasaurs such as *Bradysaurus bairdi* and *Embrithosaurus schwarzi* are quite distantly related to turtles. Slightly closer to turtles is the smaller (2 m) and more lightly built *Deltavjatia vjatkensis*. Closer still is a clade of pareiasaurs with prominent cranial bosses: *Scutosaurus karpinskii*, *Elginia mirabilis*, *Pareiasuchus nasicornis*, *P. peringueyi* and *Pareiasaurus serridens*. The nearest relatives of turtles are the late, dwarf (<1 m long) heavily armoured forms *Anthodon serrarius* and *Nanoparia pricei*. However, there is conflicting evidence suggesting that pareiasaurs are the monophyletic sister group of turtles: the shortest tree consistent with this hypothesis is only 248 steps.

The evolution of the dermal armour within pareiasaurs sheds light on how the distinct carapace of turtles might have evolved. The most primitive arrangement of dermal armour is exhibited by the early, basal pareiasaurs *Bradysaurus*, *Embrithosaurus* and *Deltavjatia* (Fig. 1). Here, the osteoderms are small isolated discs restricted to the dorsal midline, immediately above the vertebral column. The flanks, and much of the dorsum, were unarmoured. Dermal armour in pareiasaurs therefore initially did not serve a protective function. Rather, this arrangement of osteoderms has been shown to have a postural role, providing extra areas of insertion for the axial musculature¹¹. It is not

surprising that the evolution of these osteoderms coincides with a size increase which would have created support problems (Fig. 1, node D).

The next stage in the evolution of the pareiasaur armour occurred at node G and is exhibited by forms such as *Scutosaurus* and *Elginia* (Fig. 1). In these pareiasaurs, the osteoderms are larger and have spread to cover the entire dorsal region. They are mostly separate, but are suturally united over the girdles. The dermal armour in these forms clearly served a protective function (though this does not preclude other roles). In *Anthodon* and *Nanoparia*, a further elaboration occurs (Fig. 1, node H): all osteoderms are united to each other, forming a rigid covering over the entire dorsum. From this condition, the turtle condition might have evolved by fusion of this rigid osteodermal mosaic with the underlying vertebrae and ribs.

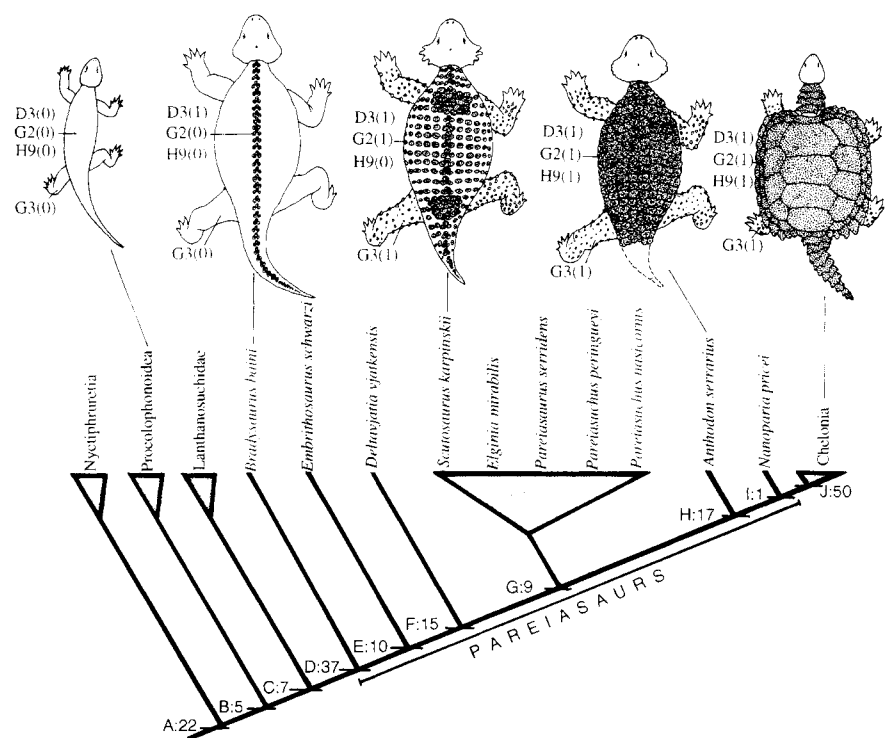
Whether pareiasaurs are monophyletic or paraphyletic with respect to turtles, the changes within pareiasaurs leading from *Bradysaurus* to *Nanoparia* provide important clues as to how the carapace in turtles might have arose. It is notable that an identical pattern of elaboration of dermal armour occurs in two other reptilian groups, archosauromorphs¹² and placodonts¹³. In these groups also, dermal armour first appears in primitive forms as a series of isolated osteoderms over the dorsal midline, and later becomes elaborated to form a rigid covering over the entire back.

Thus, in archosauromorphs, placodonts and pareiasaurs, dermal armour first arises to perform a supporting function, and only later in certain forms becomes co-opted for protection: an 'exaptation' *sensu* Gould and Vrba¹⁴. Contrary to previous, intuitive assertions¹⁵, dermal armour in turtle ancestors probably did not initially function defensively.

One of the bizarre features of the turtles is the incorporation of the dorsal vertebrae and ribs into the rigid carapace. This, of course, renders the trunk region completely inflexible, and the vertebrae have become ankylosed or even suturally united with each other and with the ribs^{1,16}. Again, evolutionary changes within pareiasaurs (Figs 1, 2) provide important clues as to how such a bizarre morphology could have evolved.

Typical primitive reptiles possessed long, slender, highly flexible trunk regions, with 25 or more presacral vertebrae^{4,17} (for example, *Owenetta*; Fig. 1). As an extant urodeles and lepidosaurs, lateral body undulations contributed significantly to each stride^{17,18}, resulting in an axial-driven¹⁹ locomotory system. Early, primitive pareiasaurs such as *Bradysaurus* (Fig. 1), *Embrithosaurus* and *Deltavjatia* would have retained some horizontal flexibility of the vertebral column. The small isolated osteoderms restricted to the dorsal midline would not have precluded such movements. However, the presacral count was reduced to twenty, and the body was short and stout (facilitating storage and fermentation of

FIG. 1 A cladogram illustrating the relationships of pareiasaurs, based on the most parsimonious tree obtained in a recent phylogenetic analysis¹⁰. Turtles are tentatively placed within pareiasaurs, near *Nanoparia pricei*. However, there remains the possibility that turtles are the sister group to a monophyletic Pareiasauridae. The number of synapomorphies diagnosing each clade is shown, although only some of these can be discussed in the following list; for a fuller discussion, see refs 9 and 10. The derived condition diagnostic of the relevant node is listed first, the primitive condition found in all other taxa is listed afterwards in parentheses. Asterisk denotes traits illustrated here and in Fig. 2. Nodes A to C: see ref. 9. Node D: D1, large size, at least 1 m long, excluding tail (small size, <50 cm excluding tail); D2, short robust body, 21 or fewer presacral vertebrae (elongated body, 24 or more presacral); *D3, dermal armour present at least over dorsal midline (dermal armour absent); for other synapomorphies, see ref. 4. Node E: E1, seven to nine cusps on each upper and lower marginal tooth (seven or fewer cusps per tooth). Node F: *F1, basal tubera positioned anteriorly, midway between basiptyergoid processes and occipital condyle (basal tubera positioned closer to occipital condyle); F2, 19 or fewer presacral vertebrae (20 or more presacral); F3, humeral torsion 45 degrees or less (torsion 60 degrees); *F4, dorsal edge of posterior coracoid oblique (dorsal edge of posterior coracoid nearly horizontal); *F5, cleithrum absent (cleithrum present in *Bradysaurus*, *Embrithosaurus*, nictiphuretians, and primitive procolophonoids, but absent in derived procolophonoids); F6, nine to eleven cusps on each upper and lower marginal tooth (nine or fewer cusps per tooth). Node G: G1, twenty or fewer caudal vertebrae (more than twenty caudals); *G2, dermal armour covers entire dorsum (dermal armour restricted to dorsal midline in *Bradysaurus*, *Embrithosaurus* and *Deltavjatia*, and absent in other taxa); *G3, limbs covered in conical bony studs (limbs without dermal armour); G4, more than eleven cusps on each lower marginal tooth (eleven or fewer cusps on each lower marginal tooth). Node H: *H1, scapula blade cylindrical (scapula blade flat and plate-like); *H2, ulnar trochlea on humerus is a discrete tubercle (ulnar trochlea is a groove); *H3, radial condyle on humerus located terminally (radial condyle located ventrally); *H4, humeral torsion 20 degrees (humeral torsion 45 degrees or more); H5, olecranon process of ulna reduced (olecranon process prominent in all other pareiasaurs, and in most related taxa, but convergently reduced in



some procolophonoids); *H6, trochanter major enlarged into a prominent tubercle (trochanter major present as a weak swelling in other pareiasaurs, absent in other taxa); *H7, trochanter major positioned very proximally, directly under capitellum (trochanter major positioned distal to capitellum in other pareiasaurs, absent in other taxa); H8, cnemial crest on tibia greatly reduced (cnemial crest is a prominent ridge); *H9, osteoderms united over entire body (osteoderms separate); H10, more than eleven cusps on each upper marginal tooth (eleven or fewer cusps on each upper marginal tooth). Node I: *I1, dorsal ribs expanded anteroposteriorly, intercostal spaces absent (ribs slender, separated by wide intercostal spaces). Reconstructions from left to right: *Owenetta* (Procolophonidea) after BPI 1/4195. *Bradysaurus bairdi* after BMNH R1971, *Scutosaurus karpinski* after PIN 2005/1532-8, *Anthodon serratus* after BPI 1/548, *Proganochelys* (Chelonina) after ref. 16. Labels refer to numbered traits discussed above: (0), primitive; (1), derived. Institutional abbreviations as in ref. 9.

ingested plant matter, and also probably inertial homeothermy²⁰). In these animals, lateral flexion of the trunk is unlikely to have contributed greatly to stride length. Retraction and long-axis rotation of the propodials¹⁸ would have been much more important. Thus, the change from the axial-driven locomotory system primitive for reptiles to the limb-driven one characteristic of turtles¹⁹ began in the earliest pareiasaurs (node D in Fig. 1).

Subsequent changes within pareiasaurs further reduced trunk flexibility. Another dorsal vertebra was lost, resulting in a pre-sacral count of only 19 (node F), the osteoderms spread over the entire dorsal region (node G) and fused with each other (node H), and the ribs widened and came to abut each other, obliterating the intercostal spaces (node I). Thus, the body of the late, dwarf pareiasaur *Nanoparia pricei* would have been almost totally inflexible. Thick, abutting ribs (Fig. 2q) have been shown to hinder lateral bending of the vertebral column²¹, and these movements would have been further restricted by the united osteoderms covering the entire dorsum. Prevention of movement at the vertebral and rib articulations would have allowed these elements to ankylose with each other and fuse with the closely overlying

osteoderms—as eventually happened in turtles. There is evidence that increasing emphasis on limb-driven locomotion accompanied the stiffening of the trunk region in pareiasaurs. The gradual reduction in humeral torsion (nodes F and H), elaboration of the elbow joint, and enlargement of the trochanter major, the insertion for one of the major femoral retractor muscles (node H), suggests that the limbs gradually increased their power and range of movement. The number of crenulations on the edges of each marginal tooth also increases (at nodes E, F, G and H), reflecting increasing adaptation and commitment to herbivory²². This is functionally linked to the other changes: loss of speed and agility in heavily armoured terrestrial reptiles appears to preclude carnivory, and only aquatic and amphibious reptiles (for example, some archosaurs, placodonts and turtles) have broken this ecological constraint. Again, whether pareiasaurs are monophyletic or paraphyletic with respect to turtles, the changes within pareiasaurs leading from *Bradysaurus* to *Nanoparia* demonstrate that evolution of the rigid body of turtles could have proceeded in small gradual steps, with functionally linked changes in several systems (axial skeleton, dermal armour, limbs, dentition) occurring in loose correlation.

The macroevolutionary model that best accounts for the above observations is 'correlated progression'⁷: the idea that evolutionary changes in one feature of the organism allow changes to take place in some other traits, changes which in turn influence the further evolution of the original feature. These feedbacks and interactions mean that the different organs appear to co-evolve with each other. This speculative application of the model as it applies to the origin of turtles from pareiasaurs will only focus on the organ systems mentioned above, though obviously the full pattern is much more intricate. For example, the initial shift to herbivory in pareiasaurs led to the development of a massive, inflexible body for storage and fermentation of ingested plant matter (Fig. 3, arrow 1). This resulted in an increased emphasis on

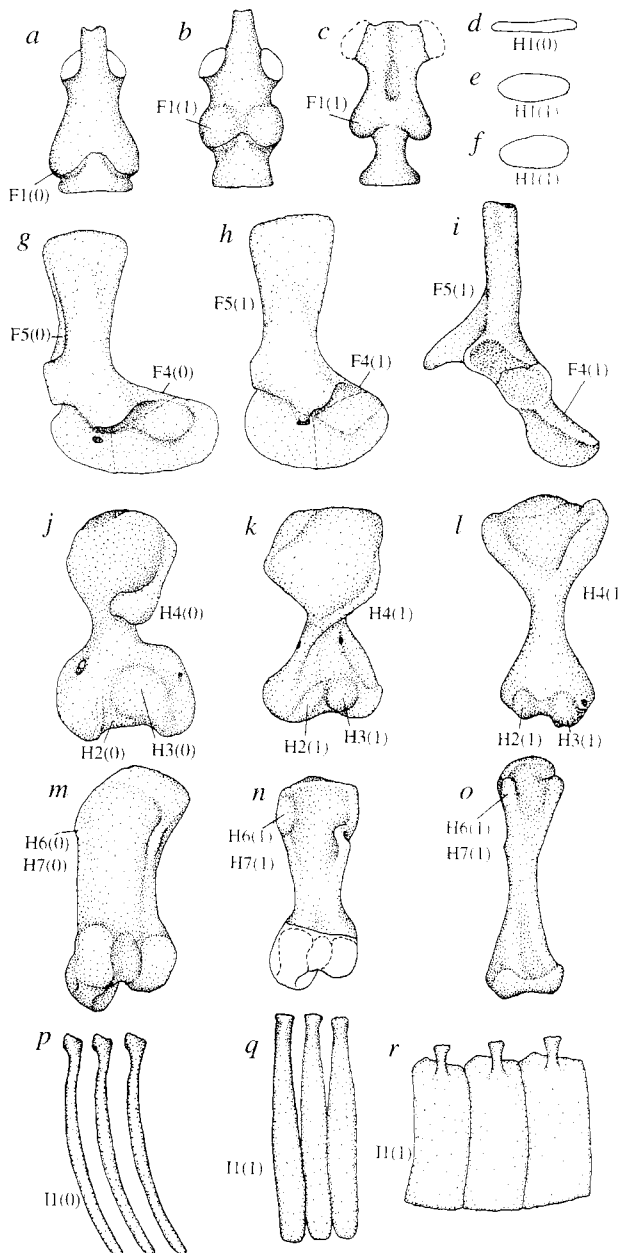


FIG. 2 Illustration of selected anatomical regions in primitive pareiasaurs, derived pareiasaurs, and turtles. Labels refer to numbered traits discussed in legend to Fig. 1. (0), primitive state; (1), derived state. Ventral view of basisphenoid and basioccipital in a, *Bradysaurus baini*; b, *Pareiasuchus peringueyi*; c, *Palaeochoersis* (Chelonia). Cross-section of scapula blades of d, *Embrithosaurus schwarzi*; e, *Anthodon serrarius*; f, *Proganochelys*. Left lateral view of scapulacoracoids of g, *Embrithosaurus schwarzi*; h, *Scutosaurus karpinskii*; i, *Proganochelys*. Ventral view of humerus of j, *Bradysaurus* sp.; k, *Anthodon serrarius*; l, *Proganochelys*. Ventral view of femur of m, *Bradysaurus baini*; n, *Anthodon serrarius*; o, *Proganochelys*. Ventral view of dorsal ribs of p, *Scutosaurus karpinskii*; q, *Nanoparia pricei*; r, *Chrysemys* (Chelonia). Parts a, m after BMNH R1971; b after BPI 1/4105; c after ref. 25; d, g after GSP CBT112; e after BPI 1/548; f, l, o after ref. 16; h, p after PIN 2005/1532; j after SAM 4350 and 6238; k after SAM 10074 and AMG 4068; n after SAM 10074; q after BPI 1/7; r after ref. 1. Institutional abbreviations as in ref. 9.

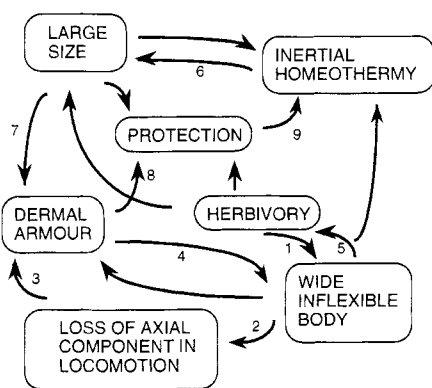


FIG. 3 Postulated selective interactions between various traits in pareiasaurs. These complex interconnections mean that changes in one trait influence the evolution of other traits, leading to a macroevolutionary pattern called 'correlated progression'¹⁷.

limb-driven locomotion (arrow 2). The reduced need for body flexibility during locomotion would have allowed greater development and consolidation of protective dermal armour (arrow 3) which in turn would have further reduced body flexibility (arrow 4). The gradual diminution of agility and speed would have locked the heavily armoured terrestrial pareiasaurs further into a herbivorous niche (arrow 5). It is not hard to see how such a system of positive feedback could have led to the gradual evolution of the totally rigid body of turtles. Similarly, the large body size in early pareiasaurs probably evolved partly in response to a shift towards inertial homeothermy²⁰ (Fig. 3, arrow 6). This in turn led to support problems alleviated by the development of dermal armour (arrow 7), which later expanded over the body and was co-opted for protection (arrow 8), which, in forming a thick dense insulating layer over the body, might have further enhanced inertial homeothermy as well (arrow 9). The shell of turtles has been shown to confer thermoregulatory advantages²³. The evolutionary changes within pareiasaurs towards 'turtle-ness' illustrate how an organism adapts to a selective regime composed not of the external surroundings alone, but of an interaction between the external surroundings and its own morphology²⁴. □

Rapid adaptive camouflage in tropical flounders

V. S. Ramachandran, C. W. Tyler, R. L. Gregory, D. Rogers-Ramachandran, S. Duensing, C. Pillsbury & C. Ramachandran

Brain and Perception Laboratory, 0109, University of California at San Diego, La Jolla, California 92093-0109, USA

DESPITE the commonly held view that flatfish can change their surface markings to match their background pattern¹, there have been few systematic studies^{2,3} and it has recently been claimed⁴ that their capacity for such adaptive changes is minimal. Here we show that the tropical flatfish *Bothus ocellatus* can achieve pattern-matching with surprising fidelity. By adjusting the contrast of different sets of 'splotches' of different grain size (or spatial frequency) on the skin, the fish can blend into a wide range of background textures in just 2–8 seconds.

Sumner² placed flatfish of the genus *Paralichthys* on checkerboard patterns and noticed that after a few days the fish seemed to develop large splotches as if to mimic the checks. These experiments were recently repeated in two genera^{4,5}, giving rise to claims that their ability to change had been grossly exaggerated and that for the most part the flatfish simply had a 'universal texture' that allowed it to blend into any environment. It was also suggested that many of the photographs of flounders showing apparent camouflage (including those in refs 2 and 3) contained artefacts produced by film nonlinearities that tended to enhance the apparent resemblance between the fish and its background.

Even without any dynamic change, the flatfish can blend surprisingly well with a wide range of variegated backgrounds⁴. But we wondered whether it was also capable of dynamic changes in pattern. Five specimens of the Caribbean shallow-water flatfish *Bothus ocellatus* were used in the experiment. They were initially housed in a tropical salt-water aquarium at 30 °C and transferred to four small 25 × 20 × 30 cm rectangular plastic tanks for testing. Their eyes are mounted on short stalks and move independently (like a chameleon's eyes) through 180°, and the fish could track and pounce on a thin vertical bar moving horizontally.

The following patterns covered the bottoms of the test tanks: a coarse gravel; printed checkerboards, either 1-cm or 2-mm checks; a homogeneous grey sheet, the luminance of which was identical to the mean luminance of the checks; or fine, yellow beach sand. While in the home tank, all fish appeared almost identical in colour and pattern to the yellowish beige gravel on the floor (similar to that in Fig. 1a). When we transferred the fish (one per tank) each changed its surface markings within 2–8 s to resemble closely its new background. Each fish was then moved to a different, randomly selected, tank and once again exhibited a strong change in pattern to match the new background. As long as the fish were healthy and active, such changes could be brought about several times (Fig. 1).

We were concerned that the resemblance between the fish and the various backgrounds might be an optical illusion: that the effect seen in Fig. 1 might be in the eye of the beholder and the fish might not have changed significantly. We therefore took colour photos of the same fish on different backgrounds and then simply cut them out and displayed them on a plain background. It was clear that the pattern on the fish was changing physically. We also took black and white half-tone pictures (Fig. 1b–d) and asked 20 naive subjects to choose the background pattern that most closely matched each of the three cut-out fish. They did this correctly in 58 of 60 trials, providing unequivocal evidence that the fish were changing physically in the appropriate direction. The reason that Saidel's results^{4,5} were not as clear may be that he used cold-water rather than tropical flatfishes and that the latter may have evolved a greater ability to engage in pattern matching. Indeed, our own

Received 3 October; accepted 8 December 1995.

1. Zangerl, R. in *Biology of the Reptilia* Vol 1 (eds Gans, C. & Bellairs, A. d'A.) 311–399 (Academic, London, 1969).
2. Rieppel, O. J. *Zool.* **231**, 487–509 (1993).
3. Pough, F. H., Heiser, J. B. & McFarland, W. N. *Vertebrate Life* 3rd edn (Macmillan, New York, 1989).
4. Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).
5. Burke, A. C. *Am. Zool.* **31**, 616–627 (1991).
6. Lee, M. S. Y. *Science* **261**, 1716–1720 (1993).
7. Kemp, T. S. *Mammal-like Reptiles and the Origin of Mammals* (Academic, London, 1982).
8. Laurin, M. & Reisz, R. R. *Zool. J. Linn. Soc.* **113**, 165–223 (1995).
9. Lee, M. S. Y. *Biol. Rev.* **70**, 459–547 (1995).
10. Lee, M. S. Y. *Zool. J. Linn. Soc.* (submitted).
11. Frey, E. in *Third Symposium on Mesozoic Terrestrial Ecosystems. Short Papers* (eds Reif, W.-E. & Westphal, F.) 93–97 (Attempto, Tübingen, 1984).
12. Sereno, P. C. *Soc. Vert. Paleont. Mem.* **2**, 1–53 (1991).
13. Peyer, B. & Kuhn-Schnyder D. *Traite de Paleontologie* **5**, 459–486 (1995).
14. Gould, S. J. & Vrba, E. S. *Paleobiology* **8**, 4–15 (1982).
15. Pritchard, P. *Acta zool. Cracov.* **31**, 625–686 (1988).
16. Gaffney, E. S. *Bull. Am. Mus. Nat. Hist.* **194**, 1–263 (1990).
17. Sumida, S. S. *Univ. Calif. Publ. Zool.* **122**, 1–133 (1990).
18. Edwards, J. L. in *Major Patterns in Vertebrate Evolution* (eds Hecht, M. K., Goody, P. C. & Hecht, B. M.) 553–577 (Plenum, New York, 1977).
19. Zug, G. R. *Misc. Publ. Mus. Zool. Univ. Michigan* **142**, 1–98 (1971).
20. King, G. M. *The Dicotyledons: A Study in Paleobiology* (Chapman and Hall, London, 1990).
21. Kemp, T. S. *J. Zool.* **191**, 193–239 (1980).
22. Hottel, N., III. *Am. Mid. Nat.* **53**, 88–114 (1955).
23. Mackay, R. S. *Nature* **204**, 355–358 (1964).
24. Lewins, R. & Lewontin, R. *The Dialectical Biologist* (Harvard University Press, Cambridge, MA, 1985).
25. Rougier, G. W., de la Fuente, M. S. & Arcucci, A. B. *Science* **268**, 855–858 (1995).

ACKNOWLEDGEMENTS. The many people and organizations acknowledged in ref. 9 are here thanked again. I thank the Australian Research Council for funding, and R. Shine and colleagues for use of laboratory and office facilities.

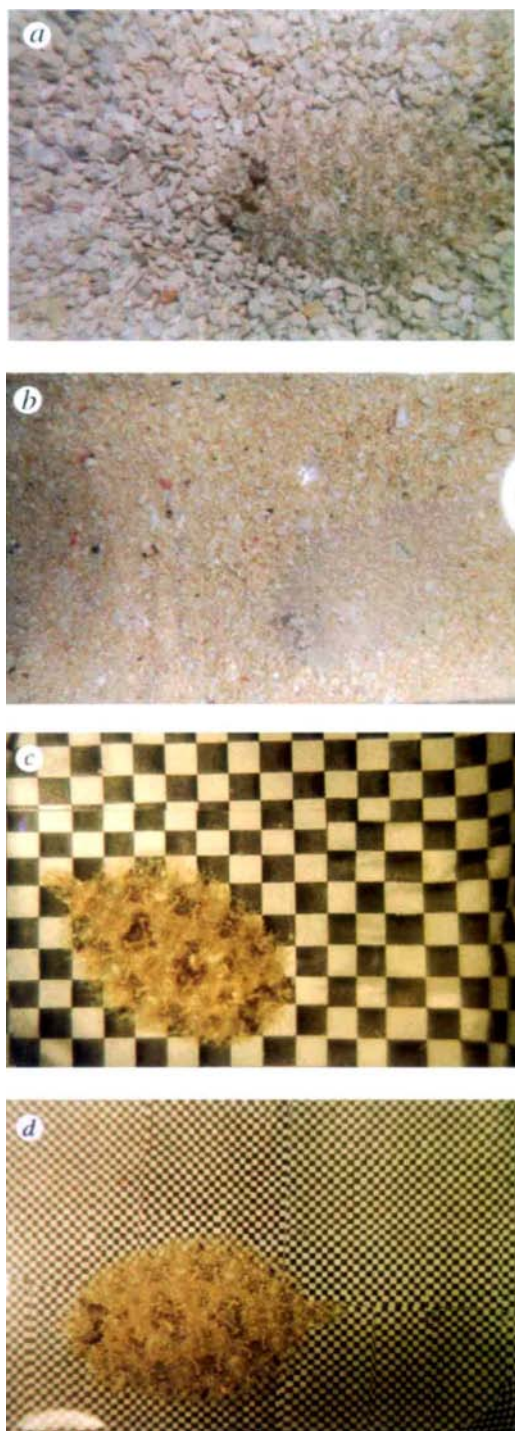


FIG. 1 Photographs of the same fish on 4 different backgrounds. The adaptive change took 2–8 s in each case. We ruled out the possibility that these effects were simply a consequence of the fish being transparent. This was a problem mainly with smaller specimens (<4 cm), but transparency in the specimens studied was also ruled out on the following grounds. First, analysis of the videotapes showed that there was a delay before the fish resembled the background. Second, the blotches on the fish had no spatial coherence with the background squares. Third, after the fish had changed its pattern, we placed an oval piece of grey opaque cardboard between it and the pattern and this did not eliminate the camouflage (photographs were also taken for subsequent analysis). And fourth, when the fish was displaced passively with a ruler, the pattern moved with the fish. The texture changes were unequivocal but emphasis should not be placed on the colour changes as these were not measured objectively (the exact tints seen in these reproductions could be artefacts produced by lighting). We also noted that as the fish were cycled through the tanks repeatedly they seemed to learn to respond faster to the background pattern. It remains to be seen whether this curious form of perceptual learning is specific to the particular pattern to which the fish is exposed.

initial results with cold-water genera (for example, *Paralichthys*) were very similar to those obtained by Saidel.

A convenient way to characterize visual texture is by means of the Fourier transform of the flounder pattern. From photographs of the fish that had had time to adapt to a variety of different background patterns, we obtained the radial power spectrum of the back of the fish, and a corresponding region of the background pattern, averaged over orientation and phase. A principal components analysis of these spectra showed that there were three predominant channels of independent information about the background texture being expressed in the flounder camouflage patterns (Fig. 2). These may be most easily described as a low-frequency/high-frequency opponent channel, a predominantly mid-frequency channel, and a channel with a narrow peak at 8 cycles per fish and a suggestion of higher harmonic structure. The important result from this analysis was that (like human colour vision) flounder patterns have three dimensions of useful control

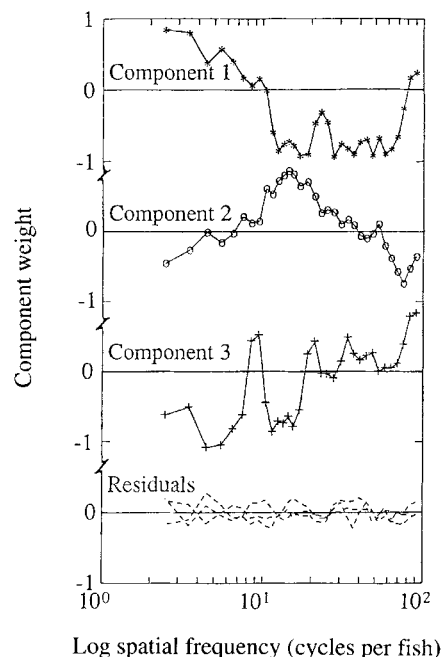


FIG. 2 Correlation analysis of the background texture information mimicked by the flounder patterns. To condense the information, we took a radial average of the 2-D power spectra around successive annular bands with widths of 1/4 octave. This provides a 1-D graph of the spatial-frequency content of the pattern of the fish or the background, averaged over orientation and phase. From photographs of the fish on 23 different background patterns to which they had had time to adapt, we obtained the radial power spectrum of the back of the fish (excluding the motile fringe and tail regions, which were semi-transparent), and an oval region of the background matching the shape of the fish. The degree of adaptive matching of the fish to the background was obtained by means of a cross-correlation matrix of the spectral power at each background frequency with that at each frequency in the flounder pattern. A principal components analysis (singular value decomposition of the cross-correlation matrix of 33×33 spatial frequencies) showed that there were three predominant channels of independent information about the background texture being expressed in the flounder camouflage patterns, as depicted by the six curves. (The eigen values of the first six components were 5.68, 5.02, 3.03, 0.94, 0.87 and 0.75.) The three significant components (full curves) may be most easily described as a low-frequency/high-frequency opponent channel, a predominantly mid-frequency channel, and a channel with a narrow peak at 8 cycles per fish and a suggestion of higher harmonic structure. Further components (dashed curves) are of low amplitude and do not appear to contain much structured information. The flounder patterns have three dimensions of useful control by the environmental information. The particular spectral form of the underlying mechanisms may consist of any linear combination of the depicted functions, which correspond to a vector space of the control parameters between the background texture and the fish patterns.

FIG. 3 Different surface markings on *Bothus ocellatus*. The fish seemed to have some degree of independent control over the different types of markings. These markings were A, large 'H'-shaped splotches lining the margins of the fish; B, small dark rings within each 'H' and elsewhere on the surface, and similar light rings scattered over the surface; C, small central black spots, one in the middle of each ring; D, stippling over the surface of the fish; E, a central 'eye spot' in the middle of the fish which was made up of an 'H' splotch with a ring and central spot in the middle; and F, a 'spectacle frame' around the eyes. Additionally, in between each ring and the central black spot was a pale area that was lighter than the yellow-brown colour of the rest of the fish. Each set of markings was essentially a cluster of specific size and shape composed of melanocytes. The luminance gradient seen across the whole photograph is artefactual.

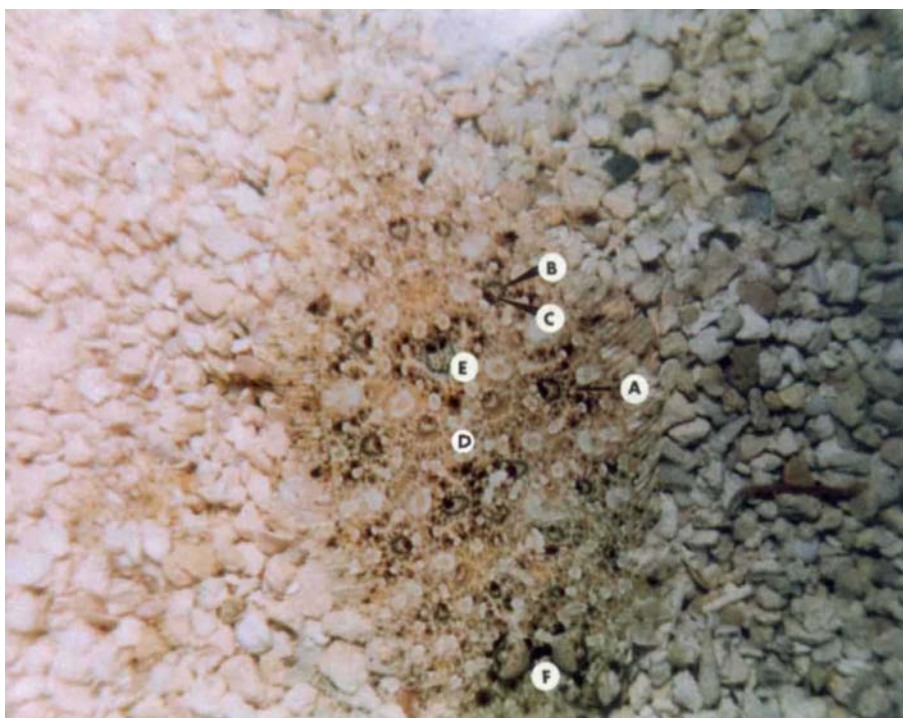
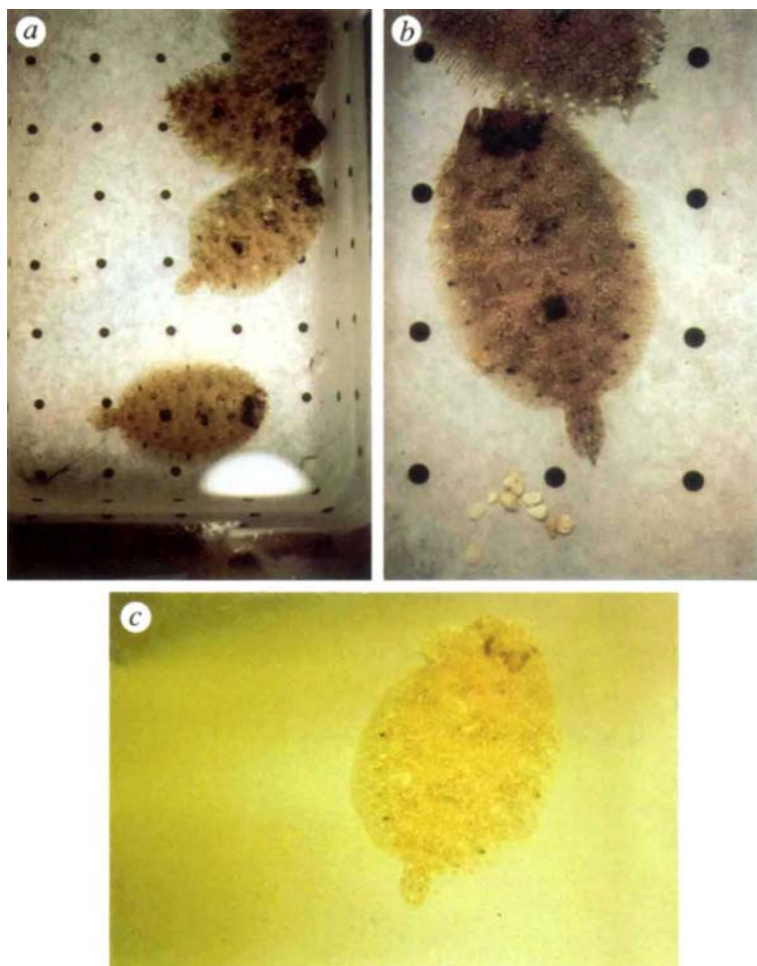


FIG. 4 Response of specimens to sparse black polka dots (a, b) and a homogeneous light grey (c). These adaptive changes took several minutes rather than seconds. In a and b, the fish increased the contrast of the central 'eye spot' and the dark spectacles around the eyes while the rest of the fish became pale. A grey background (c), the mean luminance of which was equal to this pattern, did not produce this effect. Similarly, when the fish was placed on a photonegative of a (sparse white polka dots on black) the fish became much darker and developed prominent white spots near its margins, an effect that was not seen on a homogeneous black background. Again, it is difficult to avoid the conclusion that these white spots were made more prominent to mimic the sparse white polka dots in the surround.



by the environmental information over presently unknown combinations of the six categories of elements identified in Fig. 3.

Studies under a dissecting microscope revealed that the fish has at least six categories of surface marking (Fig. 3). Hence the effector response in the skin must be organized in a modular fashion, as noted in refs 4 and 5. By adjusting the contrast of these sets of markings in different ratios, the fish could 'blend' into a wide range of natural backgrounds in much the same way that all spectral colours can be produced by mixing just three primaries in a trichromatic visual system. This argument implies, of course, that the fish must have independent visual control of each set (or subset) of markings, a possibility that requires verification. There may be 'feature detectors' in the fish visual centres that are specialized for detecting different spatial frequencies of textures in the environment and these might exert direct control over the corresponding set of marks on the skin surface. Indeed, there might be a map of effector neurons in (say) the tectum, so that focal electrical stimulation might produce selective contrast enhancement of specific spatial frequency components on the skin.

These experiments show clearly that the flounders' skin patterns were changing in an adaptive manner, but how frequent and reliable are these changes? To explore this we placed our five fish successively for 5 min on each of three types of background (large checks, small checks, and uniform grey of the same mean luminance). They were cycled through this whole sequence six times with each sequence videotaped for analysis using a Sony HiFi Handycam with a timer. Behaviour was considered to be adaptive only if the fish appeared to match the background within the first 20 s and retained the match for the remaining time on that pattern. It was important to use this rigid criterion as the fish also sometimes engaged in non-adaptive pattern changes when disturbed (for example, a blanching took place in less than a second as the fish flipped into a high-speed avoidance reaction). The results showed conclusively that all five animals were capable of adaptive behaviour; they matched their background on 91% of trials. Surprisingly, the adaptation was often completed in 2–8 s, which is considerably less time than quoted by Sumner² (several days) or Mast³ (several hours). Thus a neural, rather than humoral, mechanism must be involved.

In our last experiment we placed two fish on the patterns shown in Fig. 4. Remarkably, the fish increased the contrast of just two spots, the central 'eye spot' and the dark spectacles around the eyes, whereas the rest of the fish became pale. Thus, the fish had obviously made an attempt to match the sparse black polka dots by turning on only two black spots on the skin surface. A grey background (Fig. 4c), the mean luminance of which was equal to this pattern, did not produce the effect.

A distraction manoeuvre that we observed also deserves comment. When pursued around a large tank, the fish would disturb the sand in one location while darting back to bury itself quickly in a different position, thereby deceiving us into thinking that it was actually buried in the first location.

In summary, we conclude that tropical flatfishes are indeed capable of adaptive camouflage and that such adaptation can occur with surprising rapidity. The manner in which such precise visual control of melanophores is achieved remains to be determined. Interestingly, squids⁶ can also achieve rapid camouflage in a similar manner, a remarkable example of convergent evolution. □

Received 23 October; accepted 13 December 1995

1. Bomford, L. *Camouflage and Color*. (Dorset, New York, 1992).
2. Sumner, F. B. J. *exp. Zool.* **10**, 409–505 (1911).
3. Mast, S. O. *Bull. U.S. Bur. Fish.* **34**, 173–232 (1914).
4. Saidel, W. M. in *Sensory Biology of Aquatic Animals* (eds Atema, J., Fay, R., Popper, A. N. & Tavolga, W.) (Springer, New York, 1988).
5. Saidel, W. M. thesis, MIT, Cambridge, MA (1977).
6. Packard, A. in *Cephalopod Neurobiology* (eds Abbott, N. J., Williamson, R. & Maddock, L.) 231–267 (Oxford Univ. Press, 1995).

ACKNOWLEDGEMENTS. These experiments were originally begun by V.S.R. and R.L.G. at the Scripps Institute of Oceanography. We thank T. Bullock for help and the Biomedical Research Fund (UCSD) and NIMH for financial support.

Vasoactive effects of fibrinogen on saphenous vein

Robert C. J. Hicks, Jonathan Golledge,
Rachid Mir-Hasseine & Janet T. Powell*

Cardiovascular & Pulmonary Research Division, Charing Cross & Westminster Medical School, London W6 8RF, UK

NORMAL plasma fibrinogen concentrations are critical to haemostasis. Higher fibrinogen concentrations are associated with increasing risk of atherosclerotic disease¹ and with graft stenosis and occlusion after saphenous vein bypass surgery^{2,3}. Vein graft stenosis is characterized by the localized proliferation of intimal smooth muscle cells, causing narrowing of the graft with increased risk of thrombotic occlusion. In rabbit arteries, fibrinopeptide B is reported to have both vasoconstrictor and mitogenic properties^{4,5}. We report here that fibrinopeptides had no vasoactive effects on saphenous vein rings; however, fibrinogen (0–2 μ M) affected an endothelium-dependent relaxation, followed by recontraction at higher concentrations. The fibrinogen-mediated relaxation was inhibited by K⁺-channel blockers and antibodies to ICAM-1. Coupled signalling pathways for the synthesis of vasoactive mediators and mitogens could underlie the association between fibrinogen and the development of vein graft pathology.

We studied the changes in isometric tension of saphenous vein rings in response to thrombin, fibrinogen and fibrinopeptides. Vein rings exhibited endothelium-dependent relaxation to increasing doses of thrombin, maximum relaxation $42 \pm 4\%$ at 1 U ml^{-1} which increased to $63 \pm 5\%$ in the presence of $0.25 \mu\text{M}$ fibrinogen, but no response to fibrinopeptides (A + B) alone, at concentrations of up to $5 \mu\text{M}$. Unexpectedly, vein rings showed a concentration-dependent relaxation in response to fibrinogen (0–2 μM), this relaxation being reversed at higher concentrations of fibrinogen (Fig. 1). The response to fibrinogen was not observed when vein rings were denuded of endothelium (Fig. 1). However, when intact vein rings were incubated with fibrinogen and the supernatant transferred to a vein ring denuded of endothelium, progressive relaxation occurred (Fig. 1). Other plasma proteins, albumin and transferrin at concentrations of up to $12 \mu\text{M}$, did not cause relaxation of vein rings or attenuate the response to fibrinogen. Addition of the fibrinogen fgD.fgE fragment complex (0–15 μM) did not effect relaxation of precontracted vein rings but preincubation of vein rings with the fgD.fgE complex ($10 \mu\text{M}$) abolished the relaxation in response to fibrinogen (Fig. 2). These experiments suggest that specific fibrinogen receptors on saphenous vein endothelium mediate vasomotor responses.

Nitric oxide and prostacyclin are the two best characterized vasodilators synthesized by endothelium, with nitric oxide stimulating the synthesis of cGMP in the underlying smooth muscle⁶. However, preincubation of vein rings with indomethacin ($10 \mu\text{M}$), the nitric oxide synthase inhibitor L-NAME (0.1 mM) or the nitric oxide trap haemoglobin ($100 \mu\text{g ml}^{-1}$) only effected a small reduction of the fibrinogen-mediated relaxation (data for L-NAME and indomethacin in Fig. 1 and for haemoglobin in Fig. 2). The lipoxigenase inhibitor nordihydroguaiaretic acid ($10 \mu\text{M}$) did not alter the fibrinogen-mediated relaxation, but inhibited the vasoconstriction observed at higher fibrinogen concentrations (Fig. 1). Preincubation of vein rings with both indomethacin and L-NAME reduced the maximum relaxation in response to fibrinogen from $63 \pm 3\%$ to $40 \pm 3\%$, $P = 0.03$ (Fig. 1). The concentration of cGMP extracted from the vein rings of five patients was measured by radioimmunoassay, at baseline, after treatment with $2 \mu\text{M}$ fibrinogen or $0.5 \mu\text{M}$ A23187 (which evokes the synthesis of endothelial nitric oxide): the concentrations of cGMP were 29.9 ± 4.8 , 39.0 ± 6.4 and $116.5 \pm 31.6 \text{ fmol mg}^{-1}$ protein, respectively. The modest increase in cGMP concentration in response to fibrinogen contrasts with the fourfold increase in cGMP concen-

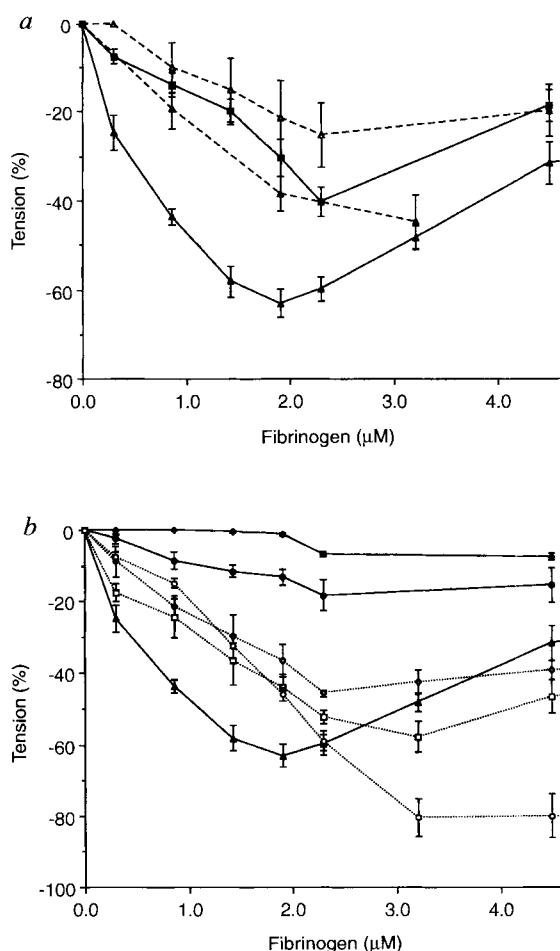
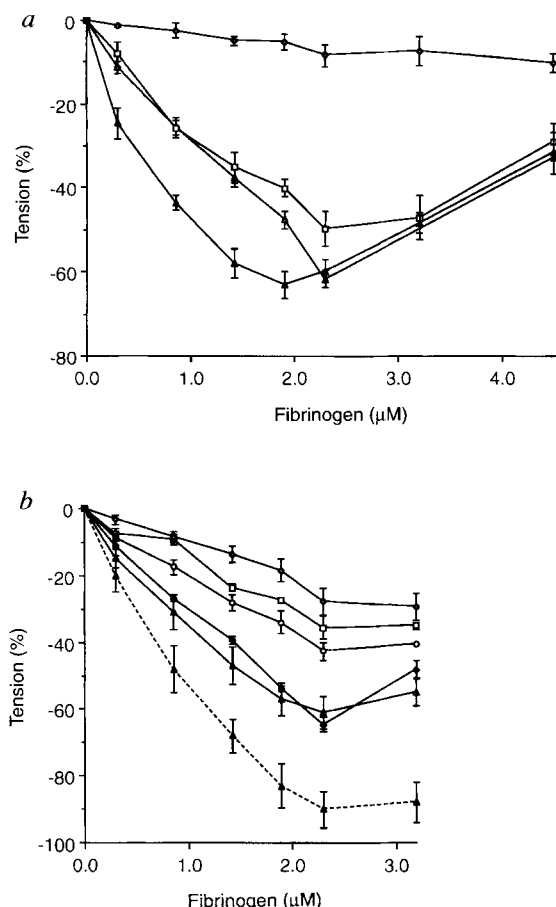


FIG. 2 Changes in isometric tension of saphenous vein rings with endothelium in response to cumulative doses of fibrinogen, after pre-incubation with proteins. **a**, Preincubation of vein rings with either $8 \mu\text{M}$ transferrin (Δ), $n = 6$, or with haemoglobin at $100 \mu\text{g ml}^{-1}$ ($\square$), $n = 4$, had little effect on the dose-response curves for fibrinogen. Preincubation of vein rings with $10 \mu\text{M}$ fibrinogen fragment fgD.fgE ($\diamond$), $n = 5$, largely abolished the response to fibrinogen. **b**, The concentration response curve to fibrinogen was not affected by preincubation of vein rings with the Fab_1 fragment of a mouse monoclonal IgG_1 to CD5 at $20 \mu\text{g ml}^{-1}$ ($\blacklozenge$), $n = 4$. Preincubation of vein rings with increasing concentrations of a Fab_1 fragment of a mouse monoclonal IgG_1 to ICAM-1 effected an increasing inhibition of the fibrinogen response: ($\circ$) $10 \mu\text{g ml}^{-1}$, $n = 4$, ($\square$) $20 \mu\text{g ml}^{-1}$, $n = 5$, ($\diamond$) $40 \mu\text{g ml}^{-1}$, $n = 3$. At a concentration of anti-ICAM-1 of $20 \mu\text{g ml}^{-1}$ the relaxation was reduced by 50% in response to cumulative doses of fibrinogen, $P = 0.002$. The response to fibrinogen was exaggerated in veins that had been exposed to a pulsatile flow circuit, mean pressure 100 mm Hg, with oxygenated Krebs solution for 45 min, conditions which increase the expression of ICAM-1 on the endothelial surface. These experiments used paired vein rings ($n = 6$); excised vein ($\blacktriangle$); vein after exposure to the arterial pressure circuit for 45 min ($\blacktriangle$, dashed line). **METHODS.** These are described in Fig. 1. The experiments where vein rings were preincubated with antibodies used paired rings from the same donor, one ring preincubated with Fab_1 fragments and the other ring preincubated with transferrin. After the dose-response curve to fibrinogen in the presence of Fab_1 fragments, the rings were washed well with Krebs solution, reequilibrated for 45 min, recontracted with phenylephrine and a further dose-response curve to fibrinogen constructed. Antibodies to ICAM-1 (CD54) were obtained from R&D Systems, Oxford, UK.

FIG. 1 Changes in isometric tension of saphenous vein rings with and without endothelium in response to cumulative doses of fibrinogen. **a**, For vein rings with endothelium ($\blacktriangle$), $n = 15$, the maximum relaxation of $63 \pm 3.1\%$ is achieved at a fibrinogen concentration of $1.8 \mu\text{M}$, followed by recontraction. Preincubation of vein rings with both L-NAME (0.1 mM) and indomethacin ($10 \mu\text{M}$) ($\blacksquare$), $n = 5$, still showed marked relaxation in response to fibrinogen, although the mean maximum relaxation was reduced to $40 \pm 3\%$. Vein rings denuded of endothelium (Δ , dashed line), $n = 4$, show minimal change in isometric tension, however, when supernatant was transferred from vein rings with endothelium, stimulated with 3 concentrations of fibrinogen, progressive relaxation occurred ($\blacktriangle$, dashed line), $n = 3$. **b**, For vein rings with endothelium preincubation with barium chloride (2 mM) ($\blacklozenge$), $n = 5$, gliclazide ($30 \mu\text{M}$) ($\bullet$), $n = 6$, inhibited the relaxation in response to increasing doses of fibrinogen. Untreated vein rings ($\blacktriangle$). Preincubation of vein rings with L-NAME (0.1 mM) ($\diamond$), $n = 5$, or indomethacin ($10 \mu\text{M}$) ($\square$), $n = 6$, alone provided some attenuation of the relaxation at low fibrinogen concentrations. Preincubation of vein rings with nordihydroguaiaretic acid ($10 \mu\text{M}$) ($\circ$), $n = 5$, also inhibited the recontraction at higher concentrations of fibrinogen.

METHODS. Saphenous vein was collected at bypass surgery from normotensive, non-diabetic subjects, who had not smoked for at least 5 years. These procedures were approved by the local ethical committee. After excision the vein was placed immediately into ice-cold Krebs solution and the rings prepared and mounted in the organ chamber¹¹. After equilibration the rings were submaximally contracted with phenylephrine ($1\text{--}10 \mu\text{M}$). Fibrinogen (from either Kabi-Pharmacia or Sigma) was dialysed against Krebs solution and the purity of the solution (at 8 mg ml^{-1}) assessed by SDS polyacrylamide gel electrophoresis under reducing conditions. Accumulative doses of fibrinogen were added at 5–8 min intervals: the response to fibrinogen developed slowly over 4–5 min and then stabilized. Vein rings were denuded of endothelium by the passage of a fine wire across the luminal surface and the efficacy of this process confirmed by scanning electron microscopy. Vein rings were preincubated in the organ chamber with drugs and proteins for 30 min before the addition of fibrinogen. At the end of each experiment the vein rings were relaxed by the addition of sodium nitroprusside ($10 \mu\text{M}$) to confirm the functional response of the smooth muscle.



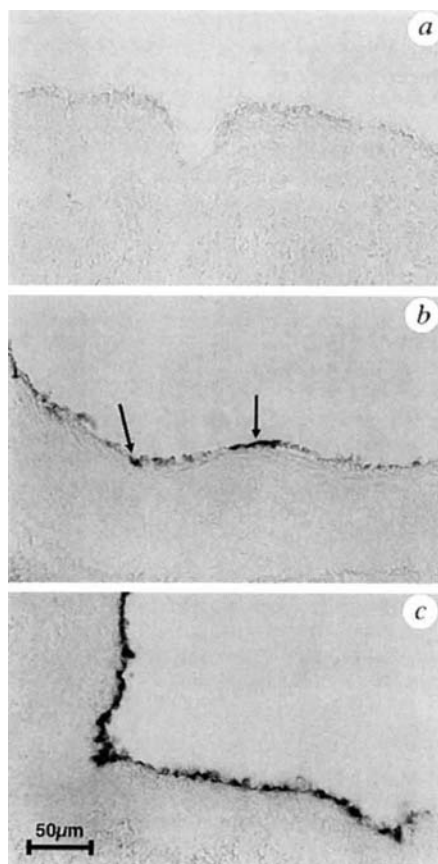


FIG. 3 ICAM-1 on saphenous vein endothelium. There is intermittent black immunostaining for ICAM-1 on the endothelial surface of freshly excised vein (b), the staining on vein exposed to pulsatile arterial pressure for 45 min is continuous and increased in intensity (c), whereas there is minimal staining on the control which was incubated with anti-CD5 as the primary antibody (a).

METHODS. Saphenous vein was fixed in Zamboni's solution for 4 h and 8- μ m sections prepared using a cryostat. Primary mouse monoclonal antibodies against ICAM-1 (Serotec, Oxford, UK) were used at 1 in 800 dilution and tissues were stained using the avidin-biotinylated complex peroxidase method.

tration in response to A23187. Such results suggest that neither nitric oxide nor prostacyclin are the principal vasodilators in the fibrinogen-mediated vasorelaxation of saphenous vein.

A third vasodilator, endothelium-derived hyperpolarizing factor, whose action has been attributed to the activation of K^+ channels, is released by arterial endothelium in response to a variety of agonists⁷. Preincubation of saphenous vein rings with ouabain (10 μ M), non-selective K^+ -channel blockers (2 mM Ba^{2+} , 3 mM tetraethylammonium) or a blocker of ATP-sensitive K^+ channels in vascular smooth muscle (30 μ M gliclazide) all inhibited the fibrinogen-mediated relaxation (data for Ba^{2+} and

gliclazide in Fig. 1). (Gliclazide did not inhibit endothelium-dependent relaxation in response to A23187.) After maximal contraction of vein rings with 40 mM KCl (achieved by isotonic replacement of NaCl), rather than phenylephrine, there was no relaxation in response to fibrinogen. Following submaximal contraction of vein rings with lower concentrations of KCl (10 mM) vasoactive effects of fibrinogen were observed again: maximum relaxation $45 \pm 8\%$ at 3 μ M fibrinogen. These results suggest that fibrinogen binds to specific receptors on saphenous vein endothelium and initiates a signalling cascade which depends on the opening of ATP-sensitive K^+ channels.

Human umbilical vein endothelial cells have receptors for fibrinogen, which appear to influence both endothelial migration and the adhesion of leukocytes^{8,9}. In particular, fibrinogen appears to stimulate, in a concentration-dependent manner, the adhesion of leukocytes to ICAM-1 on cultured human umbilical vein endothelial cells⁸. We used a Fab₁ fragment of a monoclonal antibody to human ICAM-1 to investigate further the identity of the fibrinogen receptor that signals the relaxation of saphenous vein rings. This Fab₁ fragment (0–0.4 μ M) did not have vasoactive effects, but preincubation of vein rings with this Fab₁ fragment resulted in a concentration-dependent inhibition of the relaxation in response to fibrinogen (Fig. 2). After washout of the Fab₁ fragment and reequilibration, the maximum relaxation at 2 μ M fibrinogen returned to $61 \pm 5\%$. The maximum relaxation at 2 μ M fibrinogen reduced significantly to $28 \pm 4\%$ in the presence of 40 μ g ml⁻¹ (0.8 μ M) of the Fab₁ fragment, $P = 0.01$. In contrast, at the same concentration of Fab₁ fragment the endothelium-dependent relaxation to A23187 remained unaltered. Preincubation of vein rings with control Fab₁ fragments did not influence the response of vein rings to fibrinogen (Fig. 2). In addition, preincubation of vein rings with the peptide GRGDSP (150 μ g ml⁻¹), which interrupts the reaction between fibrinogen and the endothelial integrin $\alpha_3\beta_3$ (ref. 9), did not inhibit the fibrinogen-mediated relaxation (data not shown). Endothelial expression of ICAM-1 is increased in response to shear stress¹⁰. After exposure of saphenous vein to arterial flow for 45 min, the intensity of immunostaining for ICAM-1 on endothelium increased two-to-threefold (Fig. 3) and there was a heightened sensitivity of the vasomotor response to fibrinogen: the mean maximum relaxation, at 2 μ M fibrinogen, increased from $63 \pm 5\%$ to $90 \pm 6\%$, $P = 0.03$ (Fig. 2). These experiments indicate that vasorelaxation is a consequence of fibrinogen binding to ICAM-1 on saphenous vein endothelium.

The normal concentration of fibrinogen in whole blood is about 4–6 μ M. Fibrinogen interacts with several blood components including erythrocytes, leukocytes and platelets, so that the concentration of fibrinogen at the endothelial surface may be lower. Therefore the concentration of fibrinogen that elicits vasomotor responses from saphenous vein *in vitro* lies in the physiological range. Our experiments indicate that fibrinogen binds to ICAM-1 on saphenous vein endothelium and initiates signalling pathways leading to the synthesis of vasoactive mediators, other than nitric oxide and prostacyclin. Perhaps, at higher fibrinogen concentrations, these signalling pathways also stimulate the exaggerated smooth muscle cell proliferation that results in vein graft stenosis. □

Received 4 December 1995; accepted 8 January 1996.

- Levenson, J., Giral, P., Razavian, M., Gariepy, J. & Simon, A. *Arterioscler. Thromb. vasc. Biol.* **15**, 1263–1268 (1995).
- Hicks, R. C. et al. *Eur. J. vasc. Surg.* **9**, 415–420 (1995).
- Wiseman, S. et al. *Br. Med. J.* **299**, 643–664 (1989).
- Moore, P. K., Hussaini, I. & Bhardwaj, R. J. *Pharm. Pharmacol.* **40**, 558–561 (1988).
- Singh, T. M., Kadowaki, M. H., Glagov, S. & Zarins, C. K. *Am. J. Surg.* **160**, 156–159 (1990).
- Waldman, S. A. & Murad, F. J. *cardiovasc. Pharmacol.* **12** (suppl. 5), S115–S118 (1988).
- Nagao, T. & Vanhoutte, P. M. *Am. J. Respir. Cell. molec. Biol.* **8**, 1–6 (1993).

- Dejana, E. et al. *J. clin. Invest.* **75**, 11–18 (1985).
- Languino, L. R. et al. *Cell* **73**, 1423–1434 (1993).
- Nagel, T., Atkinson, W. A., Dewey, C. F. & Gimbrone, M. A. *J. clin. Invest.* **94**, 885–891 (1994).
- Higman, D. J., Greenhalgh, R. M. & Powell, I. T. *Br. J. Surg.* **80**, 1242–1245 (1993).

ACKNOWLEDGEMENTS. We thank R. M. Greenhalgh and D. Lane for providing saphenous vein and the fibrinogen fragment [gD.fgE, respectively. Serier Laboratories for a gift of gliclazide, R. Turner for cGMP extraction and the Medical Research Council and the AFG Research Foundation for financial support.

Mapping of a susceptibility locus for Crohn's disease on chromosome 16

Jean-Pierre Hugot*, Pierre Laurent-Puig*, Corine Gower-Rousseau†, Jane M. Olson‡, John C. Lee§, Laurent Beaugerie||, Isam Naom¶, Jean-Louis Dupas#, André Van Gossum☆, Groupe d'Etude Thérapeutique des Affections Inflammatoires Digestives**, Marianne Orholm††, Catherine Bonaiti-Pellie‡‡, Jean Weissenbach§§, Christopher G. Mathew¶¶, John E. Lennard-Jones§, Antoine Cortot†, Jean-Frédéric Colombel† & Gilles Thomas*|||

* Laboratoire de Génétique des Tumeurs, INSERM U 434, Institut Curie, 75231 Paris, France

† Registre des Maladies Inflammatoires du Tube Digestif du Nord-Ouest de la France, CHRU, 59037 Lille, France

‡ Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio 44109, USA

§ St Mark's Hospital, London EC1V 2PS, UK

|| Service de Gastro-Entérologie et d'Hépatologie, Hôpital Rothschild, 75571 Paris, France

¶ Division of Medical and Molecular Genetics, UMDS Guy's & St Thomas's Hospitals, London SE1 9RT, UK

Service d'Hépatogastro-Entérologie, Hôpital Nord, 80054 Amiens, France

☆ Service de Gastro-Entérologie et d'Hépatogastro-Pancréatologie, Hôpital Erasme, 1070 Bruxelles, Belgium

** GETAID: Groupe d'Etude Thérapeutique des Affections Inflammatoires Digestives, 33 blvd de Picpus, 75012 Paris, France

†† Department of Medical Gastroenterology, Herlev Hospital, Copenhagen, Denmark

‡‡ Unité d'Epidémiologie des Cancers, Institut Gustave Roussy, 94805 Villejuif, France

§§ CNRS URA 1922 Généthron, 1 rue de l'Internationale, 91002 Evry, France

CROHN'S disease (CD) and ulcerative colitis are the major forms of chronic inflammatory bowel diseases in the western world, and occur in young adults with an estimated prevalence of more than one per thousand inhabitants¹. The causes of inflammatory bowel diseases remain unknown, but genetic epidemiology studies²⁻⁵ suggest that inherited factors may contribute in part to variation in individual susceptibility to Crohn's disease. A genome-wide search performed on two consecutive and independent panels of families with multiple affected members, using a non-parametric two-point sibling-pair linkage method, identified a putative CD-susceptibility locus on chromosome 16 ($P < 0.01$ for each panel). The localization was centred around loci D16S409 and D16S419 by using multipoint sibpair analysis (ref. 6, and J.M.O., manuscript submitted) ($P < 1.5 \times 10^{-5}$). This region of the genome contains candidate genes which may be relevant to the pathogenic mechanism of inflammatory bowel diseases.

An initial panel of 25 caucasian families, each with at least two affected siblings, was studied (Table 1). Family members were genotyped for 270 highly polymorphic markers, with a mean heterozygosity of 0.76. Markers were selected that spread along the autosomes and the pseudo-autosomal regions, spanning approximately 37 morgans⁷, with a mean sex-averaged genetic distance between two consecutive marker loci of 12.8 cM. Only 18 intervals were greater than 20 cM, and none reached 30 cM.

Using the parameters derived from segregation analyses on northwestern European populations, which indicated a monogenic recessive mode of inheritance for CD^{5,8}, linkage analyses were initially conducted with the parametric lod score method⁹ as previously described¹⁰. Simulated estimates of power¹¹, obtained using a polymorphic marker with an heterozygosity of 0.75, at a distance of 10 cM from the disease locus, indicated that a lod score greater than 3 was expected with a probability higher than 97%. However, the maximum two-point lod score obtained experimentally was 2.04 for D16S409 (recombination fraction 0.18). Four

flanking markers located 9–42 cM from D16S409 did not reveal higher values. In contrast, lod scores providing evidence of linkage exclusion (lod score less than -2) were observed for these markers and the remaining 265 loci tested, resulting in an exclusion map of the whole genome. Taken together, these results suggest that the initial assumption derived from segregation analyses of a monogenic recessive model for CD might be an oversimplification.

In the absence of additional information on the contribution of genetic factors to CD, the data were reanalysed using a non-parametric method based on identity by descent (i.b.d.) in affected sibling pairs (sibpairs)¹². Using a one-sided test for mean proportion of alleles shared i.b.d. (greater than 0.5 for a linked marker), 4 of the 270 markers demonstrated an increased degree of allele sharing among affected sibpairs with a P -value less than 0.01 (Table 2). The four markers indicated two different genetic regions which might contain CD-susceptibility loci. Nevertheless, because of the large number of tests in such a genome-wide approach, this finding required confirmation on another panel of families.

A second independent panel of 53 families (Table 1) was genotyped for the four markers selected from the initial screen. On this second panel, two markers (D16S408 and D16S409) demonstrated a mean proportion of alleles shared i.b.d. greater than 50% with a level of significance equal or close to $P = 0.01$ (Table 2).

Additional markers mapped to a 40-cM region containing D16S408 and D16S409 were typed on the pooled family panels. Seven out of eight adjacent loci demonstrated evidence of linkage to the CD locus (Table 3). A multipoint analysis based on affected sibpair allele sharing (ref. 6 and manuscript submitted) demonstrated linkage¹³ between D16S409 and D16S419 ($P < 1.5 \times 10^{-5}$; Fig. 1).

This result enables the provisional assignment of linkage to chromosome 16 of a CD-susceptibility locus, which we call *IBD1* (for inflammatory bowel disease locus 1). At the location of the maximum lod score, the estimated ratio of the dominance genetic variance to the total genetic variance was 0.88 (manuscript submitted) and the estimated proportions of siblings sharing 0, 1 and 2 alleles i.b.d. were 0.19, 0.40 and 0.42, respectively. Taken together,

TABLE 1 Number of families and affected sibpairs in CD family panels

	Panel 1		Panel 2	
	Affected siblings per family	Equivalent number of affected sibpairs	Number of families	Equivalent number of affected sibpairs
2	19	19	40	40
3		4.33	9	19.5
4	3	10.24	2	6.83
5*	0	0	2	4.33
6	0	0	0	0
7	1	7.35	0	0
Total	25	40.92	53	70.66

Patients were classified as affected if they met the criteria for definite or probable CD, according to previously validated procedures²¹. For the analyses using the lod score method, family members were coded as unaffected if neither chronic inflammatory nor gastroenterological symptoms were found in their individual records. Relatives diagnosed as possible CD²¹, or those whose collected data were insufficient, were classified as unknown for CD status. Mixed families in which some members are affected with CD and others with ulcerative colitis were excluded. The first family panel contained families with only one generation involved; the second panel was a mixture of families with one or more generations involved. For the non-parametric linkage analyses, data were taken from affected sibpairs only. In a sibship of size, $s > 2$, $s(s-1)/2$ pairs can be formed, but they are not independent. To extract all of the pair data available from these sibships, we applied a coefficient that correlates the quantity of information contained in the s affected siblings in a family to that contained in independent pairs²². This correction provides the equivalent number of affected sibpairs given in column 3 and 5.

* Sibpairs were provided from two generations.

||| To whom correspondence should be addressed.

these observations suggest that *IBD1* may be inherited in a recessive manner. Using a method described previously¹⁴, the relative risk ratio of CD due to *IBD1* in siblings of CD patients was estimated to be 1.3, which is comparable to values proposed for the *IDDM2* and *IDDM4* loci of insulin-dependent diabetes mellitus^{12,15}.

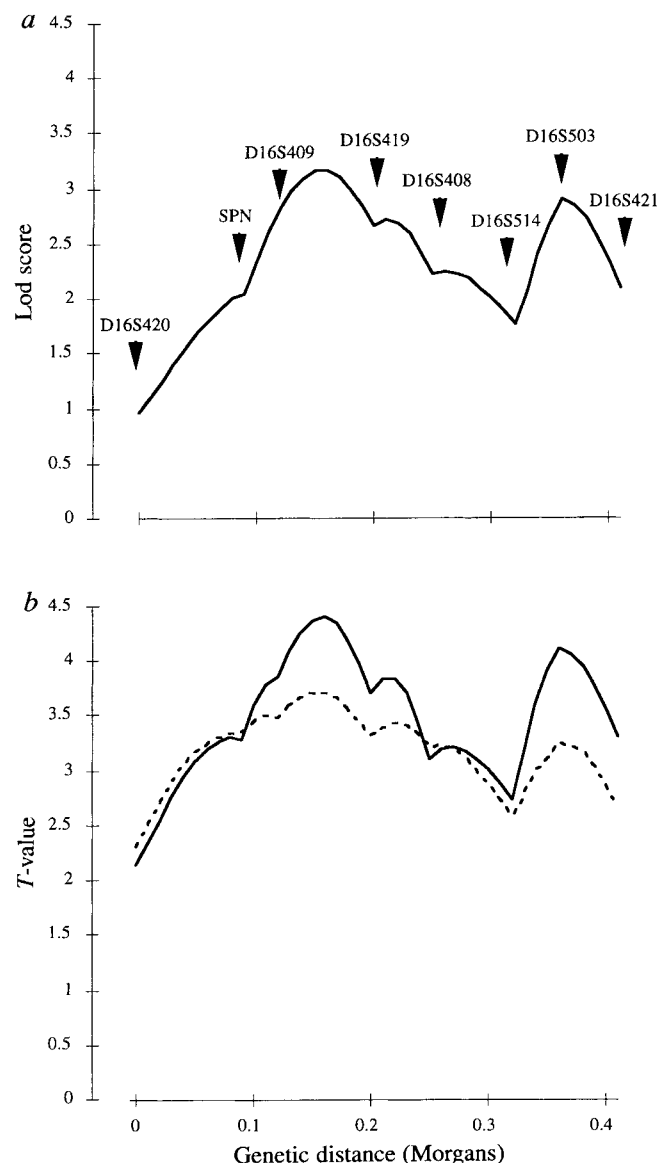


FIG. 1 Multipoint interval analysis based on sibpair allele sharing for loci in the pericentromeric region of chromosome 16. Genotype data from the loci listed in Table 3 were used to perform a multipoint analysis. *a*, The lod score; *b*, two pseudo-likelihood tests ($T\beta$, broken line; T_{z2} , continuous line). The largest lod score ($z = 3.17$) occurs 4 cM from D16S409. The largest values of $T\beta$ and T_{z2} occur 3 cM and 4 cM from D16S409, respectively ($T\beta = 3.72$, $P = 0.0001$; $T_{z2} = 4.42$, $P < 1.5 \times 10^{-5}$; both one-sided tests).

METHODS. Genetic distances between the markers were computed from the data collected on our set of families with Crohn's disease using the CRIMAP program²³. For each point of the map, marker allele sharing for each affected sibling pair was estimated from all eight markers and the likelihood maximized using the program MAPMAKER/SIBS²⁶. As the lod score cannot be assigned a P -value, owing to the inclusion of all affected sibling pairs in each family, two pseudo-likelihood tests²⁵ analogous to non-parametric tests²⁷ were developed. The first test, $T\beta$, tests the total genetic effect and is analogous to the non-parametric mean test²⁷. The second test, T_{z2} , tests the dominance effect and is analogous to the test of the proportion of pairs sharing two alleles i.b.d.²⁷; such tests are more powerful than tests of mean i.b.d. sharing for recessive diseases (ref. 27 and manuscript submitted).

IBD1 lies in the pericentromeric region of chromosome 16. Several candidate genes within 10 cM of the location of the maximum lod score may be relevant to the etiopathogenic mechanism of CD¹⁶. The most conspicuous examples are: the CD11 integrin cluster, which includes complement receptor type 3, involved in mycobacterial cell adhesion¹⁷, CD19, which is involved in B-lymphocyte function¹⁸; sialophorin, which is involved in leukocyte adhesion¹⁹; and the interleukin (IL)-4 receptor, as IL-4-mediated regulation of mononuclear phagocyte effector functions is altered in inflammatory bowel diseases²⁰.

The initial screen identified a possible CD-susceptibility locus close to D1S236 on chromosome 1. However, the linkage result for this marker was not significant ($P > 0.05$) in the second family panel (Table 2). This observation could reflect a type-1 statistical error during the first screen or heterogeneity between the two family panels. In this respect, it is notable that, in the first family panel, the affected CD members were present in a single generation, whereas the second panel was a mixture of families with affected members in one or more generations. However, attempts to demonstrate genetic heterogeneity for D1S236 in these two types of families, although suggestive, failed to reach statistical significance (data not shown).

Typing of additional CD families is required to provide an accurate description of the genetics of CD. It would better enable refinement of the location of the putative gene on chromosome 16, the further investigation of the possible involvement of a locus

TABLE 2 Assignment of a CD-susceptibility locus to chromosome 16

Family panel 1			
Locus	Mean proportion of alleles shared i.b.d.	T-test value	P value
D1S236	0.76	3.23	0.0006
D16S419	0.67	2.34	0.01
D16S409	0.67	2.38	0.01
D16S408	0.73	3.11	0.001
Family panel 2			
Locus	Mean proportion of alleles shared i.b.d.	T-test value	P value
D1S236	0.58	1.32	> 0.05
D16S419	0.53	0.58	> 0.05
D16S409	0.64	2.49	0.01
D16S408	0.64	2.29	0.011

Family panel 1 was typed for 270 loci (list of markers available from the authors on request). Only the four markers shown exhibited a positive test for linkage during the course of the initial screen with a P -value less than 0.01 ($t > 2.32$). For family panel 2, only these four markers were typed in the second panel of families. Genotyping data were obtained and verified as described previously¹⁰. I.b.d. sharing for affected siblings was deduced from parental genotype data. Whenever possible, in the case of lack of data from one parent, the corresponding genotype was reconstructed from the data provided by the unaffected members of the sibship, if present. The mean proportion of alleles shared i.b.d. was estimated from counts of the number of informative allele pairs (N) and the number that were i.b.d. (I). Data were provided only by heterozygous parents. When both parents were heterozygous and the parental origin of the alleles could be determined, each affected sibpair contributed two informative allele pairs (one from each parent). When one parent was homozygous, only alleles from the heterozygous parent were taken into account; thus each affected sibpair contributed only one allele pair. When parents were both heterozygous but exhibited identical genotype, the affected sibpair contributed only in the case of unambiguous parental transmission. This procedure means that we do not need to know the allele frequencies in the population. The mean proportion of alleles shared i.b.d. for a locus was estimated as I/N , and the statistic $T = 2(I - N/2)/N^{1/2}$ was used to test linkage¹². The value of the statistic was compared with the table of the standardized normal distribution to obtain a one-sided significance level. The estimated mean proportion of alleles shared i.b.d., the value of the test statistic T , and its degree of significance (P value) are shown.

TABLE 3 Two-point affected sibpair analysis for loci in the pericentromeric region of chromosome 16

Locus	Mean proportion of alleles shared i.b.d.	T-test value	P value
D16S420	0.55	1.72	0.044
SPN*	0.58	2.36	0.0099
D16S409†	0.58	3.41	0.0004
D16S419†	0.58	3.05	0.0014
D16S408†	0.57	3.09	0.0012
D16S514	0.55	1.69	0.0471
D16S503	0.58	2.44	0.0082
D16S421	0.51	0.3	> 0.05

Polymorphic loci from the pericentromeric region of chromosome 16, where the putative CD locus lies, were selected using information from the Genome Data Base (Johns Hopkins University, Baltimore, Maryland) and the integrated chromosome-16 map of the Center for Human Genome Studies at Los Alamos National Laboratory. These markers were typed on the two panels of CD families. The loci are listed according to the order along chromosome 16 (ref. 16) and that provided by the analysis of the data generated on the merged panels of CD families using CRIMAP²³. The data generated from the pedigree founders of the merged family panels enabled the estimation of allele frequencies for each markers that were used to compute the allele-sharing estimates and *T*-statistics using the genetic analysis program SIBPAL²⁴.

* SPN, Sialoporphin or CD43 (ref. 19).

† Markers used in the initial genome-wide search.

on chromosome 1p, and the search for additional susceptibility genes. The locus on chromosome 16 accounts for only a small fraction of the tenfold increased risk for first-degree relatives of CD patients⁴. Some of the genetic factors involved in CD may also contribute susceptibility to ulcerative colitis. Indeed, CD and ulcerative colitis share the same ethnic predisposition³, and mixed families in which some members are affected with CD and others with ulcerative colitis are commonly found⁴. This hypothesis can now be tested with markers identified in this study. □

Received 2 November; accepted 29 December 1995.

- Calkins, B. M. & Mendeloff, A. I. *Epidem. Rev.* **8**, 60–91 (1986).
- Tysk, C., Lindberg, E., Järnerot, G. & Floderus-Myrhed, B. *Gut* **29**, 990–996 (1988).
- Sonnenberg, A. & Wasseman, A. H. *Gastroenterology* **101**, 122–130 (1991).
- Orholm, M. et al. *New Engl. J. Med.* **324**, 84–88 (1991).
- Küster, W., Pascoe, L., Purmann, J., Funk, S. & Majewski, F. *Am. J. med. Genet.* **32**, 105–108 (1989).
- Kruglyak, L. & Lander, E. S. *Am. J. hum. Genet.* **57**, 439–454 (1995).
- Gyapay, G. et al. *Nature Genet.* **7**, 246–339 (1994).
- Orholm, M. et al. *Br. med. J.* **306**, 20–24 (1993).
- Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. *Am. J. hum. Genet.* **37**, 482–498 (1985).
- Hugot, J. P. et al. *Am. J. med. Genet.* **52**, 207–213 (1994).
- Ott, J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4175–4178 (1989).
- Hashimoto, L. et al. *Nature* **371**, 161–164 (1994).
- Lander, E. S. & Schork, N. J. *Science* **265**, 2037–2048 (1994).
- Risch, N. *Am. J. hum. Genet.* **40**, 1–14 (1987).
- Davies, J. L. et al. *Nature* **371**, 130–136 (1994).
- Doggett, N. A. et al. *Nature* **377** suppl. 335–365 (1995).
- Hetland, G. & Wiker, H. G. *Immunology* **82**, 445–449 (1994).
- Maloney, M. D. & Ringwood, C. A. *J. exp. Med.* **180**, 191–201 (1994).
- Manjunath, N., Johnson, R. S., Staunton, D. E., Pasqualini, R. & Ardman, B. J. *Immun.* **151**, 1528–1534 (1993).
- Schreiber, S. et al. *Gastroenterology* **108**, 21–33 (1995).
- Gower-Rousseau, C. et al. *Gut* **35**, 1433–1438 (1994).
- Hodge, S. E. *Genet. Epidemiol.* **1**, 109–122 (1984).
- Green, P. *Cytogenet. cell. Genet.* **59**, 122–124 (1992).
- Statistical Analysis for Genetic Epidemiology, release 2.2. (Dept Biometry and Genetics, LSU Medical Center, New Orleans, 1994).
- White, H. *Econometrica* **50**, 1–25 (1982).
- Rish, N. *Am. J. hum. Genet.* **46**, 229–253 (1990).
- Blackwelder, W. C. & Elston, R. C. *Genet. Epidemiol.* **2**, 85–87 (1985).

ACKNOWLEDGEMENTS. We thank the patients and their families who provided blood samples for this study; J. Belaiche, M. C. Boutron, C. Cann, J. P. Cezard, B. Duclos, T. Georgelin, V. G. Levy, V. Li, E. René and M. Stablo for family recruitment; many British physicians for providing clinical data; R. Modigliani and J. P. Gendre for support; and S. Olschwang for help with experimental procedures. This work was supported by grants from the Association Apetit, the Fonds de Recherche de la Société Nationale Française de Gastro-Entérologie, the Association Française contre les Myopathies, the Ministère de la Recherche et de l'Action Humanitaire, INSERM, CHRU-Lille, the Ferring Company, the Crohn's in Childhood Research Association, the National Association for Crohn's and Crohn's Disease (UK), NCI and NCCR. J.P.H. received a fellowship from the Ligue Nationale Française Contre le Cancer. Some of the results in this paper were obtained by using the program SAGE, which is supported by the USPHS division of Research Resources.

Functional redundancy of the muscle-specific transcription factors Myf5 and myogenin

Yukang Wang, Patrick N. J. Schnegelsberg, Jessica Dausman & Rudolf Jaenisch

Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

THE myogenic basic helix–loop–helix transcription factors, Myf5, MyoD, myogenin and MRF4, play key roles in skeletal muscle development^{1,2}. All of them induce myogenic differentiation in cultured non-muscle cells, suggesting that they might be functionally redundant. But the genes are expressed at different times during embryogenesis^{3–6} and mice carrying a mutation in any of the genes have different phenotypes^{7–13}. A rib cage defect was observed in Myf5-deficient mice, which die perinatally⁷. We investigated whether the rib cage defect was due to the failure of the early activation of the gene or to the unique interactions of Myf5 with specific downstream targets. For this we inserted a myogenin complementary DNA into the Myf5 locus by homologous recombination which simultaneously disrupted Myf5 function. We report here that mice homozygous for this myogenin gene knock-in (ki) developed a normal rib cage and were viable, therefore demonstrating functional redundancy of Myf5 and myogenin for rib formation.

The targeting vector (Fig. 1a) deleted the coding sequences in exon 1 of Myf5 encoding the basic helix–loop–helix (bHLH) domain responsible for DNA binding, dimerization and activation of myogenesis¹⁴. These sequences were replaced with the complete myogenin coding region (*myg*) followed by a stop codon. A DNA fragment containing *PGK-neo* flanked by two *IoxP* sites was inserted downstream of the *myg* segment. The targeting vector was transfected into embryonic stem (ES) cells, and three recombinant ES cell clones carrying the replacement of the endogenous Myf5 gene were identified. To remove the *PGK-neo* cassette, one of the ES cell clones was transiently transfected with the Cre recombinase expression plasmid *pMC-Cre*¹⁵ (Fig. 1a, b). Therefore, the resulting mutant allele, *Myf5^{myg-ki}*, contains the myogenin coding sequence under the control of the Myf5 regulatory elements, the exon–intron organization being unaltered. We anticipated that the mutant *Myf5^{myg-ki}* messenger RNA, its coding sequence in exon 1 being replaced by *myg* and a *IoxP* site, would be transcribed at the time of normal Myf5 activation and produce functional myogenin protein but no Myf5 protein, because of the stop codon immediately following the myogenin coding sequences. Indeed, a construct that carries the exact exon and intron sequences of the *Myf5^{myg-ki}* locus produced myogenin in COS cells of the same size as the endogenous myogenin in C2C12 myocytes. (Fig. 1d)

Three independent ES cell clones (clones C8.1–3, shown in Fig. 1b) carrying the *Myf5^{myg-ki}* allele were transmitted through the mouse germ line (Fig. 1c). To verify that the myogenin cDNA was expressed under Myf5 transcriptional control, we used reverse transcriptase polymerase chain reaction (RT-PCR) to detect Myf5 and myogenin RNAs in various embryos at different developmental stages. Figure 2a shows that, as expected, wild-type and heterozygous mutant embryos produced a transcript detected by Myf5-specific primers that was not seen in the homozygous mutant embryos (*Myf5^{myg-ki/myg-ki}*). In contrast, the *Myf5^{myg-ki/myg-ki}* mutant embryos produced no Myf5 transcript but instead a transcript detected by primers specific for *Myf5^{myg-ki}* RNA, which was also detected in the heterozygous embryos. To determine whether the *Myf5^{myg-ki}* allele activation was coincident with wild-type Myf5 allele activation, RNAs from heterozygous embryos at E8, E8.5 and E9.5 (day of gestation) were analysed. Figure 2b shows that

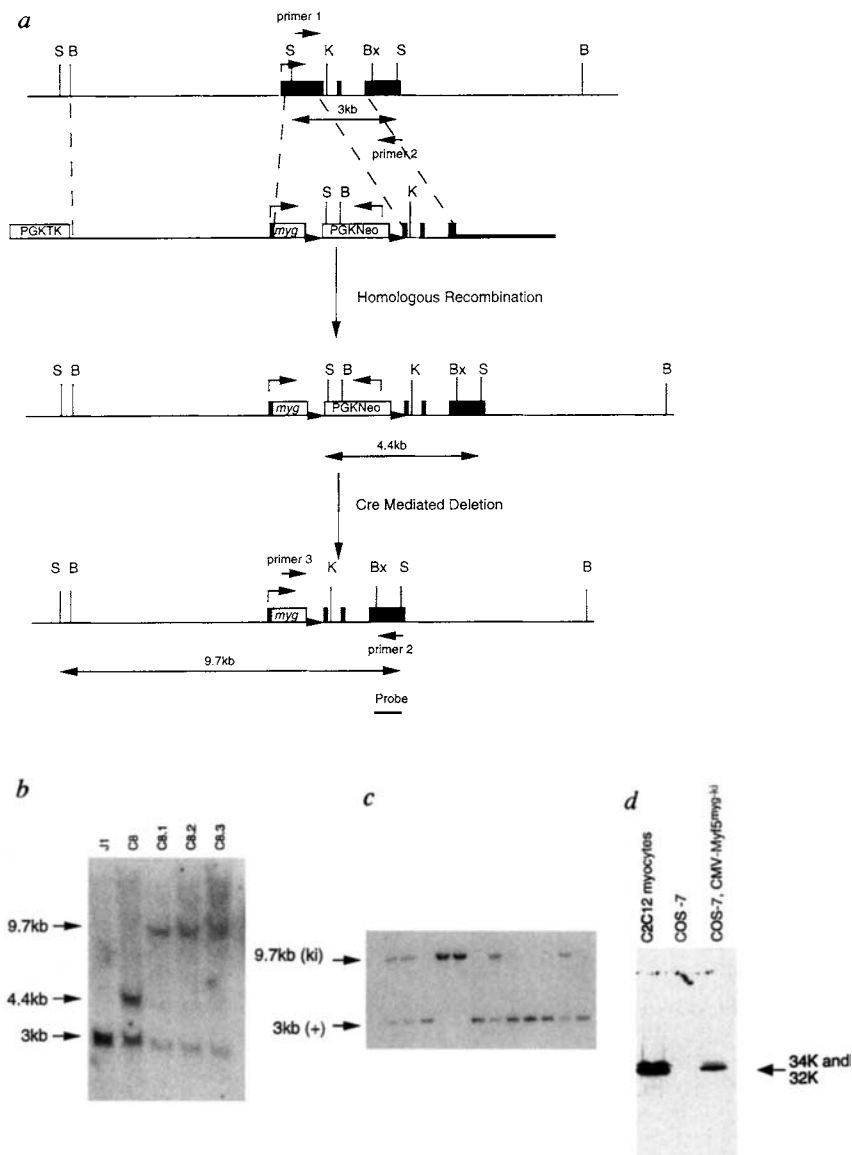
Myf5-specific RNA appeared as expected at E8, whereas myogenin RNA was only apparent at E9.5. In contrast, *Myf5^{myg-ki}* RNA was already expressed at E8 and increased at E8.5 and E9.5 with kinetics roughly paralleling that of *Myf5* RNA. These results demonstrate that *Myf5^{myg-ki}* RNA was under the transcriptional control of the *Myf5* regulatory elements and was activated before the endogenous myogenin RNA. Myotome formation has previously been shown to be delayed in *Myf5* mutant embryos. We therefore investigated whether the *Myf5^{myg-ki}* allele would rescue the delayed appearance of skeletal muscle-specific markers. Figure 2c shows that several skeletal muscle-specific genes were expressed in E9.5 *Myf5^{myg-ki/myg-ki}*, but not in *Myf5* $-/-$ embryos. The expression of those genes appeared not to depend on the activation of MyoD because it is not detected at this stage. We conclude that the *Myf5^{myg-ki}* allele rescues the myotomal-maturation defect in *Myf5* $-/-$ embryos.

In contrast to the previously generated *Myf5* $-/-$ mice, which are unable to breathe and die immediately after birth because of a rib cage defect⁷, the homozygous *Myf5^{myg-ki/myg-ki}* mice had normal rib cages, which were indistinguishable from those of wild-type

animals (Fig. 3). *Myf5^{myg-ki/myg-ki}* mice were viable, fertile, and no obvious defects in the skeletal and muscular systems have been detected. This suggests that myogenin, when expressed at the time of *Myf5* activation, fully replaces the role of *Myf5* in normal rib cage formation. A similar approach in a recent study¹⁶ was used to demonstrate the functional redundancy between *En-1* and *En-2*.

Both *Myf5* and myogenin are members of the myogenic bHLH family and have similar DNA-binding domains and can convert fibroblasts into muscle cells *in vitro*. However, they have different transcriptional activation domains. Moreover, *Myf5* and myogenin have different temporal expression patterns during embryogenesis, with *Myf5* and myogenin being expressed at E8.0 and E8.5, respectively^{3,4}. Although *Myf5*-null mice exhibit normal skeletal muscle development at birth because of the expression of MyoD¹⁷, they show a 2-day delay in myotome maturation and it has been suggested that this delay of early myotome maturation is responsible for the rib cage defect⁷. Possibly, early myogenic cells provide a permissive environment required for normal proliferation of rib precursors¹⁸. Because myogenin, expressed in the early myotomal cells under the control of *Myf5* regulatory elements,

FIG. 1 Gene replacement of *Myf5* with *myg* using the 'Cre-loxP' system. **a**, Lines 1–4 are, from top to bottom: 1, The *Myf5* locus with three exons shown as black boxes. 2, Linearized *Myf5* knock-in (*Myf5^{myg-ki}*) vector. The transcriptional directions for both *myg* and *neo* are shown. The two arrowheads represent the loxP sites. The vector sequence is shown in the thicker line. 3, The structure of the targeted *Myf5* locus after Cre-mediated deletion. The primers for detecting either *Myf5* or *Myf5^{myg-ki}* transcript are indicated. B, BamHI; Bx, BstXI; K, KpnI; S, Scal. **b**, Southern blot of ES cell clones. DNA samples were digested with Scal, and then analysed with the probe shown in **a**. J1, Parental ES clone; C8 contains *myogenin* and *PGK-neo* cassette substituted at the *Myf5* locus. Clones C8.1–3 are derived from C8 and have *PGK-neo* deleted from *Myf5* locus. **c**, The genotype analysis of the offsprings from a *Myf5* $(+/-myg-ki) \times Myf5$ $(+/-myg-ki)$ cross. The 3 kb and 9.7 kb Scal–Scal fragments represent the wild-type and *Myf5^{myg-ki}* alleles, respectively. **d**, Western analysis of myogenin. The exon 1 of the *Myf5^{myg-ki}* allele was amplified by PCR from a *Myf5^{myg-ki}* homozygous animal and then ligated to a *Myf5* genomic fragment containing the rest of the exons and introns. This *Myf5^{myg-ki}* construct was under the control of the CMV promoter. Two specific bands were seen in lanes loaded with extracts from CMV-*Myf5^{myg-ki}*-transfected COS cells, as well as C2C12 myocytes. No specific band was seen in untransfected COS cells. **METHODS**. J1 ES cells were cultured and electroporated, and the mutant mice were generated as described¹⁵. Cre-mediated gene deletion was as described¹⁵. Western blot used anti-myogenin antibody F5D²⁰.



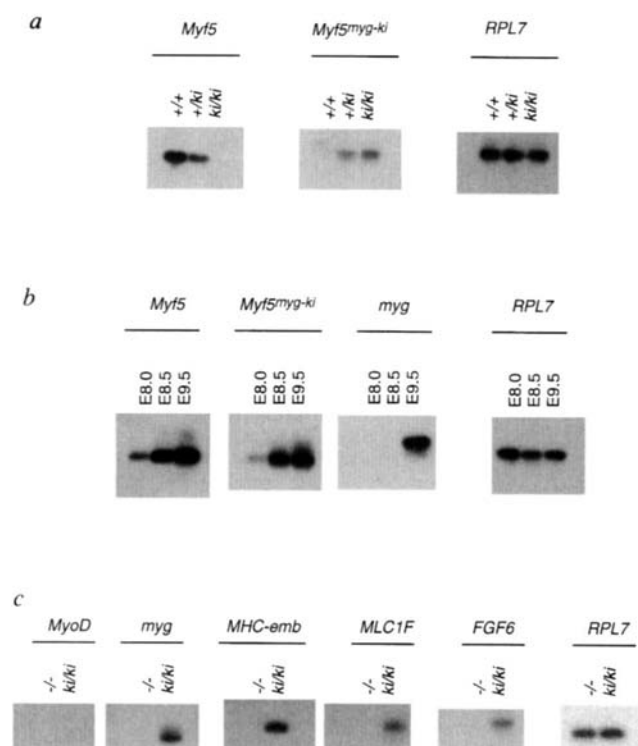
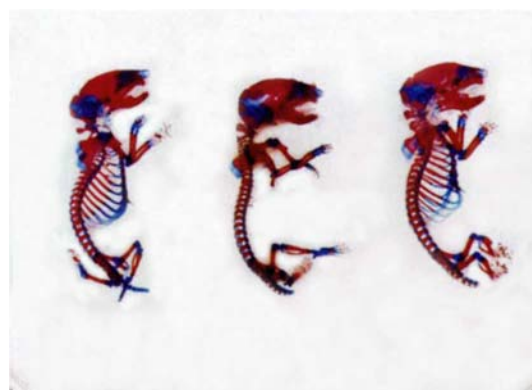


FIG. 2 RT-PCR analysis of muscle-specific gene expression in early mouse embryos. **a**, RT-PCR of *Myf5* (+/+), *Myf5* (+/*myg-ki*) and *Myf5* (*myg-ki*/*myg-ki*) embryos. Primer 1, 5' CCTGCAAAGCTTGAAGAGGA 3', from exon 1 of *Myf5*; primer 2, 5' TGGACAAGCAATCCAAGCTGG 3', from exon 3 of *Myf5*; primer 3, 5' ACAGCGCCTCCTGCAGTCCGG 3', from exon 1 of *myogenin*. Primers 1 and 2 were used for detection of the wild-type *Myf5* transcript (*Myf5*), and primers 2 and 3 were designed for detection of the transcript from the *Myf5*^{*myg-ki*} allele. Primers 1–3 are shown in Fig. 1. The primers for ribosomal protein L7 (RPL7) RNA have been described previously²¹. *Ki* and + represent *Myf5*^{*myg-ki*} and *Myf5* wild-type alleles, respectively. **b**, RT-PCR of *Myf5* (+/*myg-ki*) embryos. E8.0, 8.5 and 9.5 embryos were collected and RT-PCR was as described in **a**. The primers for detection of *Myf5* and *Myf5*^{*myg-ki*} transcripts were described in **a**. The primers for detection of endogenous myogenin transcript are: primer 4, 5' AAGCGCAGGCTCAAGAAAGT 3', from exon 1 of *myg*; and primer 5, 5' ATGCACACCCAGCTGACAGA 3', from the 3' untranslated region located in the exon 3 of *myg*. Primers 4 and 5 specifically detect endogenous myogenin transcript. **c**, RT-PCR of *Myf5* (–/–) and *Myf5* (*myg-ki*/*myg-ki*) embryos. E9.5 embryos were collected and RT-PCR was done as described in **a**. The primers for detection of *MHC-emb* and *MLC1F* and *FGF6* and *MyoD* were described previously¹³. Blots were hybridized with either *Myf5*- or myogenin- or other skeletal muscle-specific gene probes.

METHODS. Embryos were dissected from pregnant mice, and quantitative RT-PCR for muscle specific genes was essentially as described¹³. *RPL7* was used as a control to normalize input RNA and RT-PCR efficiency.



Myf5: +/+ +/- *myg-ki*/*myg-ki*

FIG. 3 Staining of newborn mouse skeletons with Alizarin red and Alcian blue. Lateral views of whole mouse skeletons to show bone (red) and cartilage (blue). The skeleton on the left is from wild-type, the middle skeleton is from a *Myf5*-null mouse and displays the rib cage defect, and the right skeleton is from a *Myf5*^{*myg-ki*}/*myg-ki* mouse and shows rescue of the rib cage defect, which was seen in every *Myf5*^{*myg-ki*}/*myg-ki* mouse analysed. METHODS. Newborn mice were eviscerated and skinned. The skins were used for genotyping, and the corpses were processed for staining as described¹².

rescues the *Myf5* mutant phenotype, we conclude that the proteins encoded by *Myf5* and *myg* are functionally redundant for normal rib development. Our results suggest that the different activation time of these two proteins during embryogenesis, rather than their different downstream targets, may be sufficient to explain the

different phenotypes of *Myf5*- and *myg* mutant mice. Nevertheless, our results do not reveal whether myogenin, expressed from the *Myf5*^{*myg-ki*} allele, is also functional in later myoblast differentiation. The interbreeding between *myg* mutants and *Myf5*^{*myg-ki*}/*myg-ki* mutants will address this question. □

Received 25 October 1995; accepted 8 January 1996.

- Olson, E. N. & Klein, W. H. *Genes Dev.* **8**, 1–8 (1994).
- Rudnicki, M. A. & Jaenisch, R. *Bioessays* **17**, 203–209 (1995).
- Ott, M.-O., Bober, E., Lyons, G., Arnold, H. H. & Buckingham, M. *Development* **111**, 1097–1107 (1991).
- Sassoon, D. et al. *Nature* **341**, 303–307 (1989).
- Hinterberger, T., Sassoon, D., Rhodes, S. & Konecny, S. *Dev. Biol.* **147**, 144–156 (1991).
- Bober, E. et al. *J. Cell Biol.* **113**, 1255–1265 (1991).
- Braun, T., Rudnicki, M. A., Arnold, H. H. & Jaenisch, R. *Cell* **71**, 369–382 (1992).
- Rudnicki, M. A., Braun, T., Hinuma, S. & Jaenisch, R. *Cell* **71**, 383–390 (1992).
- Hasty, P. et al. *Nature* **363**, 501–506 (1993).
- Nabeshima, Y. et al. *Nature* **364**, 532–535 (1993).
- Braun, T. & Arnold, H. H. *EMBO J.* **14**, 1176–1186 (1995).
- Zhang, W., Boerger, R. R. & Olson, E. N. *Genes Dev.* **9**, 1388–1399 (1995).
- Patapoutian, A. et al. *Development* **121**, 3347–3358 (1995).
- Weintraub, H. et al. *Science* **251**, 761–766 (1991).

- Gu, H., Zou, Y. R. & Rajewsky, K. *Cell* **78**, 1155–1164 (1993).
- Hanks, M., Wurst, W., Anson-Carwright, L., Auerbach, A. B. & Joyner, A. L. *Science* **269**, 679–682 (1995).
- Rudnicki, M. A. et al. *Cell* **75**, 1241–1244 (1993).
- Tajbakhsh, S. & Buckingham, M. E. *Proc. natn. Acad. Sci. U.S.A.* **91**, 747–751 (1994).
- Li, E., Bestor, T. H. & Jaenisch, R. *Cell* **69**, 915–926 (1992).
- Wright, W. E., Binder, M. & Funk, W. *Molec. cell. Biol.* **11**, 4104–4110 (1991).
- Hollenberg, S. M., Cheng, P. E. & Weintraub, H. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8028–8032 (1993).

ACKNOWLEDGEMENTS. We thank R. Curry for technical assistance and C. Beard for advice on some of the experiments. We thank C. Costigan and D. Talbot for critical reading of this manuscript. We also thank E. N. Olson for providing *myogenin* cDNA clone pMyo8, K. Rajewsky for pMC-Cre, and W. Wright for myogenin monoclonal antibody (F5D). This work was supported by an NIH/NCI grant to R.J.Y.V. was supported by a fellowship from the Muscular Dystrophy Association, and P.N.J.S. was supported by the Markey Biomedical Scientist Training Program.

Normal spatial learning despite regional inhibition of LTP in mice lacking Thy-1

M. Nosten-Bertrand^{*†}, M. L. Errington[‡],
K. P. S. J. Murphy[‡], Y. Tokugawa[§], E. Barboni^{*||},
E. Kozlova^{*}, D. Michalovich^{||}, R. G. M. Morris[#],
J. Silver[§], C. L. Stewart[☆], T. V. P. Bliss[‡]
& R. J. Morris^{*††}

Departments of ^{*}Neurobiology and [‡]Neurophysiology, National Institute for Medical Research, Mill Hill, London, NW7 1AA, UK

[§] Division of Molecular Medicine, North Shore University Hospital, Cornell University Medical College, New York 11030, USA

^{||} Department of Human Biopathology, University of Rome 'La Sapienza', Rome, 00100, Italy

[¶] Department of Cell Biology and Genetics, Erasmus University, Rotterdam 3000DR, The Netherlands

[#] Centre for Neuroscience, University of Edinburgh, Edinburgh, EH8 9LE, UK

[☆] Roche Institute of Molecular Biology, Nutley, New Jersey 07110, USA

^{††} Department of Experimental Pathology, UMDS Guy's Hospital, London Bridge, London, SE1 9RT, UK

The process of learning involves stable changes in synaptic efficacy for which long-term potentiation (LTP)¹ provides a widely adopted mammalian model. Synaptic modification induced by learning or LTP may involve the action of cell adhesion molecules²⁻⁴. One such candidate is the ubiquitous neuronal glycoprotein Thy-1 (refs 5, 6). In mice in which the gene encoding Thy-1 has been inactivated, we find a regionally selective impairment of LTP *in vivo* in the hippocampal formation: LTP is normal in area CA1 but strongly inhibited in the dentate gyrus. Spatial learning by Thy-1-deficient mice, as assessed in the watermaze⁷, is unimpaired. Thus LTP in the cortical input to the dentate gyrus seems not to be required for spatial learning.

The *Thy-1* gene was inactivated in 129/Sv embryonic stem cells by insertion of a *neo* cassette to disrupt the codon for the third amino acid of the mature protein (Fig. 1a). The inactivated gene was identified by Southern blotting (Fig. 1b), and no Thy-1 protein could be detected in *Thy-1*^{-/-} mice by immunoblotting (Fig. 1c) or immunohistochemistry (Fig. 2b, d); reduced levels of Thy-1 were present in the brains of *Thy-1*^{+/-} mice (Fig. 1c).

Thy-1-deficient mice did not differ from their littermates in breeding, behaviour or health, and no neuroanatomical differences were observed. We illustrate the latter with reference to the dentate gyrus/CA1 region of the hippocampus. The absence of Thy-1 in *-/-* mice contrasts with its abundant distribution in *+/+* mice (Fig. 2a-d). Timm's staining (Fig. 2e, f) revealed no differences in the stratified arrangement of terminal fields in the hippocampus, and no differences were detected using immunohistochemical markers of axons (for example, GAP43 and neurofilament proteins (not shown)). Similarly, the distribution of GABAergic inhibitory interneurons (Fig. 2g, h) and the number and affinity of NMDA-type glutamate receptors (Fig. 1d, e) did not differ. Preliminary data suggest that plasticity of ocular dominance during development is normal in these mice⁸.

We measured LTP *in vivo* and *in vitro* in two hippocampal

FIG. 1 Targeted disruption of the mouse *Thy-1* gene and its molecular analysis. a, Scheme of *Thy-1* genomic DNA, targeting vector, and disrupted gene. Neo, neomycin-resistance gene; TK, herpes virus thymidine kinase gene; B, *Bam*HI; Bs, *Bst*Ell; A, *Apal*; E, *Eco*RI; X, *Xho*I, b, Southern blot analysis of *Bam*HI-digested genomic DNA from F2 mice obtained by breeding the *Thy-1* null line to 129/Sv mice, probed with *Apal* fragment shown in a; the 12-kilobase (kb) fragment present in the normal *Thy-1* gene (and at half intensity in *+/-* DNA) is replaced in *-/-* mice by a 1.8 kb band; the 0.8 kb band serves as a loading marker. c, Immunoblot of brain homogenates, using polyclonal rabbit anti-Thy-1 antibodies. d, e, NMDA (N-methyl-D-aspartate)-receptor binding properties determined by ³H-MK801 binding (filled symbols, Thy-1 *-/-* membranes; open symbols, Thy-1 *+/+* membranes) to d, hippocampal membranes (inset, Scatchard plot of saturation isotherm; *K*_{obs} 2.8 nM, *B*_{max} 0.6 pmol mg⁻¹) and to e, cortical membranes (inset: % enhancement of ³H-MK801 binding to cortical membranes by L-glutamate (●, ○) and glycine (▲, △)). Points are the means of 3 independent determinations performed in duplicate. d.p.m., Disintegrations per min; M, molarity.

METHODS. The *Thy-1* gene was inactivated by inserting the 1.9-kb *Bst*Ell fragment from PGKneo cassette into the *Bst*Ell site of the third exon of *Thy-1*, disrupting the codon for the third amino acid of the mature Thy-1 protein, using homologous recombination in 129/Sv embryonic stem cells²⁵. Chimaeric mice produced were mated with C57BL/6 mice, their offspring interbred to obtain *Thy-1*^{-/-} mice that were then mated with 129/Sv mice to set up the F₂ segregating generation, on which the analysis here was based. Southern and immunoblotting followed standard procedures. ³H-MK801 binding in the presence of 100 μM L-glutamate and glycine to frozen/thawed hippocampal and cortical membranes was determined as described²⁶.

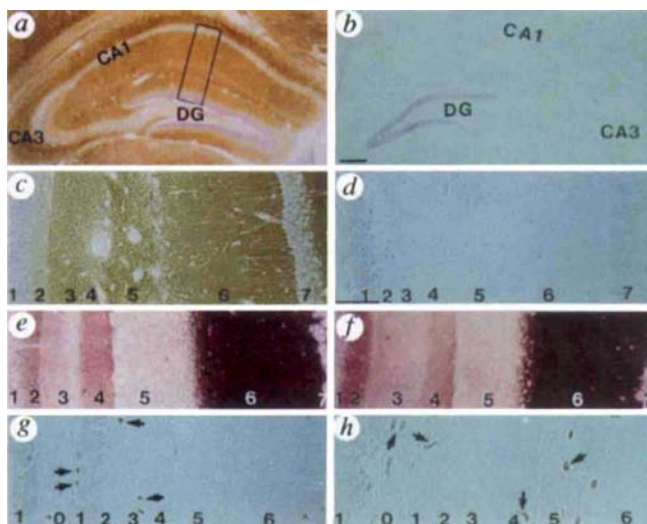
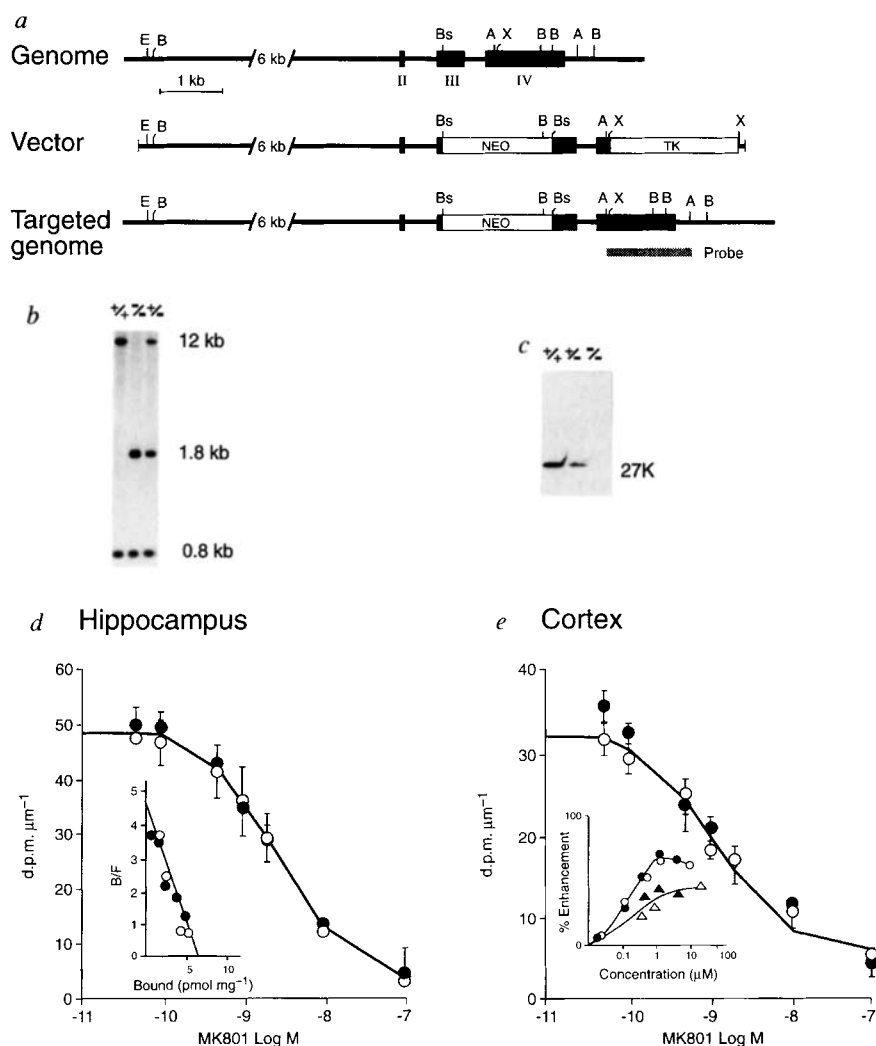


FIG. 2 Normal hippocampal cytoarchitecture in Thy-1-deficient mice. Coronal sections from *Thy-1*^{+/-} (a, c, e, g) and *Thy-1*^{-/-} (b, d, f, h) mice, immunolabelled for Thy-1 (a-d), histochemically labelled by Timm's method (e, f), or immunolabelled for GABA (g, h); nuclei throughout are mildly counterstained blue with thionin. The appropriate laminar arrangement of the main neuronal cell types, visible at low power (a, b; DG, dentate gyrus; CA1 and CA3, the respective pyramidal cell layers) is shown in greater detail for the region (outlined in a) in which LTP was studied using markers that differentiate the strata in which different axons terminate (c-f) or that identify the main, GABAergic inhibitory interneurons²⁷ (g, h; arrows indicate some of the labelled cells). The strata are: 1, granule-cell bodies; 2, dentate inner molecular layer (where axons from the hilar pyramidal cells terminate);

[†] Present address: UMB C9923 au CNRS, Hôpital de la Pitié Salpêtrière, 75013 Paris, France.

^{††} To whom correspondence should be addressed.



3, dentate mid-molecular layer (receiving medial entorhinal axons); 4, dentate outer molecular layer (receiving lateral entorhinal axons); 5, stratum lacunosum-moleculare (outer region of CA1 dendrites where entorhinal axons terminate); 6, stratum radiatum (receiving Schaffer collateral axons onto CA1 dendrites); 7, CA1 pyramidal-cell bodies; 0 (g, h only) hilar region (rich in GABAergic interneurons that innervate the dentate molecular region). Quantification of GABAergic neurons showed 31 ± 3 , compared with 29 ± 1 , cells mm^{-2} in coronal sections of dentate molecular layer, and 340 ± 16 , compared with 331 ± 25 , cells mm^{-2} in the hilar/granule-cell layers, for *Thy-1*^{+/+} and *Thy-1*^{-/-} mice respectively. **METHODS.** Mice (4–8 of each genotype with each fixative) were fixed by intraventricular perfusion with 0.5% glutaraldehyde/0.5% paraformaldehyde (Thy-1), 6% glutaraldehyde/1% paraformaldehyde/1% sodium metabisulphite ($\text{Na}_2\text{S}_2\text{O}_5$) (GABA) or 1% Na_2S in 100 mM NaH_2PO_4 (Timm's). Serial sections across the hippocampus were collected in the 3 cardinal planes and sections at 50- μm intervals were labelled with antibodies to Thy-1 (ref. 28) to GABA (Affiniti Research Products), or with Timm's stain. For quantification of the density of GABAergic cells, in 8 immunolabelled slides per brain for two brains of each genotype, the area of the dentate molecular layer (both supra- and infra-pyramidal) and hilar/granular-cell region, and labelled cells therein, was measured using BioVision software (ImproVision Ltd).

pathways where its induction is NMDA receptor-dependent¹: the perforant path input to granule cells of the dentate gyrus and the Schaffer-commissural input to pyramidal cells of area CA1. LTP of the population excitatory postsynaptic potential (e.p.s.p.) in the Schaffer-commissural projection was virtually identical in control and mutant mice, both *in vivo* (Fig. 3a) and *in vitro* (not shown). Similarly there was no difference in LTP of the population spike height, assessed *in vitro*, between the two groups (not shown). In the dentate gyrus, in contrast, we found no LTP *in vivo* in *Thy-1*^{-/-} mice, using any of eight different stimulus protocols tried. The method that produced the optimal potentiation in control mice resulted in LTP of the synaptic response (e.p.s.p.) in 9/12 control mice, and 0/14 *Thy-1*-deficient mice (Fig. 3b, e, f). LTP of the population spike was found in 11/12 control, and 0/14 *Thy-1*-deficient mice (Fig. 3e). We also investigated LTP in the dentate gyrus *in vitro*. LTP of the e.p.s.p. (Fig. 3c) and of the population spike (not shown) was similar in control and mutant mice. To obtain LTP *in vitro* in the dentate gyrus in rats, it is necessary to include a GABA_A antagonist in the perfusion medium⁹, and we confirmed that this is also the case in

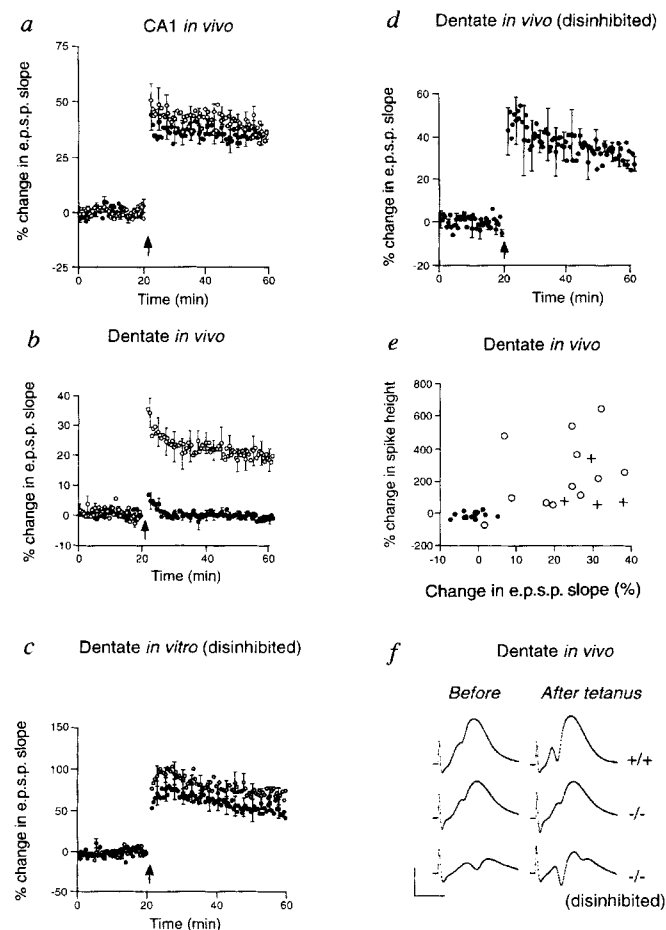
control and mutant mice (not shown). Given the ability of the disinhibited dentate gyrus to exhibit LTP *in vitro*, we wondered whether LTP *in vivo* could be rescued by including the GABA_A antagonist bicuculline in the recording pipette to produce local disinhibition¹⁰. Under these conditions, LTP was obtained in 4/4

experiments (Fig. 3d, e), establishing that suppression of LTP in the dentate gyrus of *Thy-1* mutant mice is due to enhanced inhibition.

Lesion studies have indicated that the integrity of the perforant path projection to the dentate gyrus is necessary for spatial

FIG. 3 LTP *in vivo* in *Thy-1*^{-/-} mice is normal in area CA1, but is absent in the dentate gyrus unless rescued by disinhibition using a GABA_A antagonist. **a, b, d**, Slope of the population e.p.s.p. evoked by test stimuli delivered at 30-s intervals, normalized and averaged for each *Thy-1*^{+/+} (○) and *Thy-1*^{-/-} (●) animal, and similarly for the *in vitro* experiment in dentate gyrus (**c**). LTP in each experiment was assessed as the change in response measured 35–40 min after tetanus and expressed as a percentage of the mean of the 20 responses obtained in the 10 min preceding the tetanus. Application of the tetanus is indicated by an arrow. **a**, LTP *in vivo* in area CA1, measured as the mean percentage change in slope of the e.p.s.p. in the control and mutant groups, was $34.5 \pm 4.4\%$ ($n = 8$) and $34.0 \pm 2.6\%$ ($n = 8$), respectively. In slice preparations (not shown) LTP of the e.p.s.p. was $76.4 \pm 0.07\%$ ($n = 7$) and $75.4 \pm 9.20\%$ ($n = 8$), and of the amplitude of the population spike was $304.5 \pm 57.0\%$ ($n = 6$) and $299.1 \pm 38.2\%$ ($n = 7$) for control and mutant groups, respectively. **b**, Effect of tetanic stimulation of the perforant path on the slope of the e.p.s.p. in the dentate gyrus of control and mutant mice *in vivo*. Absolute values before the tetanus (mV ms^{-1}) were 3.0 ± 0.3 (control) and 3.4 ± 0.3 (mutant mice). The mean change in the slope of the e.p.s.p. was $21.6 \pm 3.2\%$ for the control group ($P < 0.001$, *t*-test) and $-1.6 \pm 0.9\%$ (NS) for the mutant group. Corresponding values for the changes in amplitude of the population spike (not shown) were $240.1 \pm 63.3\%$ ($P < 0.001$) and $-9.2 \pm 5.1\%$ (NS). **c**, LTP in dentate gyrus *in vitro*. Mean potentiation of the slope of the e.p.s.p. was $70.2 \pm 9.5\%$ (controls, $n = 9$) and $48.5 \pm 10.4\%$ (mutants, $n = 7$) (difference between groups not significant). Inclusion of the selective NMDA-receptor antagonist D-AP5 ($25 \mu\text{M}$) in the bath blocked LTP (results not shown: mean change in e.p.s.p. slope was $-6.4 \pm 5.4\%$ (controls, $n = 4$ slices, 2 animals) and $-4.4 \pm 11.1\%$ (mutants, $n = 3$ slices, 2 animals)). **d**, Inclusion of bicuculline (10 mM) in the recording pipette enables LTP to be elicited *in vivo* in the dentate gyrus of *Thy-1*^{-/-} mice. Mean slope of the e.p.s.p. is plotted ($n = 4$). **e**, LTP in the dentate gyrus *in vivo*, plotted as the mean change in amplitude of the population spike against the corresponding change in slope of the e.p.s.p. for each animal in the control (○), mutant (●) and mutant with bicuculline (+) groups. We considered LTP to be a change of more than 10% in the slope of the e.p.s.p., or a change of more than 50% in the amplitude of the population spike (smaller changes could not be reliably distinguished from drift over time). By these criteria, LTP was exhibited by 11/12 control mice, 0/12 mutant mice, and 4/4 disinhibited mutant mice. **f**, Responses obtained from a representative animal from each group. Each trace is the average of five consecutive responses obtained immediately before and 40-min after the tetanus. Calibration: 5 ms, 5 mV.

METHODS. *In vivo* recording: Male mice, 12–15 weeks old, were anaesthetized with urethane (1.5 g kg^{-1} , intraperitoneally) and held in a semi-stereotaxic apparatus. For experiments on the dentate gyrus, a glass recording pipette was inserted 2.0 mm posterior and 1.9 mm lateral to bregma. A bipolar stimulating electrode (Rhodes SNE 100) was inserted 3.0 mm lateral to lambda, and advanced into the angular bundle to activate fibres of the perforant path. The depths of the two electrodes were adjusted to produce maximal responses in the cell-body layer. Constant-current stimuli ($50 \mu\text{s}$ duration, intensity 70 – $120 \mu\text{A}$) were delivered at 30-s intervals, and the intensity was adjusted to produce a negative population spike with an amplitude of 1 – 2 mV . The conditioning tetanus consisted of 6-trains of 6 stimuli at 400 Hz , repeated at 20-s intervals, and at double the intensity of test stimulation. No LTP was induced in the dentate gyrus of mutant mice *in vivo* with seven additional stimulation procedures, tested on a minimum of 4 animals in each case (not shown). To investigate LTP in area CA1, the recording electrode was positioned 2.0 mm posterior and 2.0 mm lateral to bregma on one side, with the stimulating electrode at the same coordinates on the contralateral side. Electrode depths were adjusted to maximize the synaptic response in stratum radiatum. A prominent population spike could usually not be recorded at this site except at high stimulus intensities which tended to produce seizure activity, and measurements were made only of the slope of the population e.p.s.p. The intensity of test stimulation was adjusted to give an e.p.s.p. with a slope $\sim 50\%$ of the



maximum obtainable. Tetanic stimulation consisted of 3 trains of 100 stimuli at 100 Hz and at test intensity, with an intertrain interval of 20 s. *In vitro* recording: Transverse $350 \mu\text{m}$ hippocampal slices were maintained in an interface chamber at 24°C as described²⁹; for slices of the dentate gyrus, the bath contained 4 mM Ca^{2+} and $100 \mu\text{M Mg}^{2+}$ and $100 \mu\text{M picrotoxin}$ ³⁰, the medial perforant path was stimulated with two trains of 20 pulses at 100 Hz , with a 30 s interval at twice test intensity. Test shocks were set to the threshold of the population spike. Recordings were made in the molecular layer. To measure LTP in the Schaffer-commissural projection to CA1 *in vitro*, area CA3 was excised. Two slices from each animal were transferred to the chamber and recordings made from the cell body layer of one slice to measure the population spike, and from stratum radiatum of the other to measure the population e.p.s.p. Monopolar tungsten stimulating electrodes (A-M systems) were placed in stratum radiatum of each slice to activate Schaffer commissural fibres. Test shocks were applied at 30-s intervals, at an intensity that evoked a response that was roughly 25% of the initial maximum response. The conditioning tetanus consisted of 3 trains of 100 shocks at 100 Hz , at 10 s intervals. Statistical analysis: significance of within-group changes produced by tetanic stimulation was assessed by paired 2-tailed *t*-tests, and differences between groups by unpaired 2-tailed *t*-tests. All experiments were done blind.

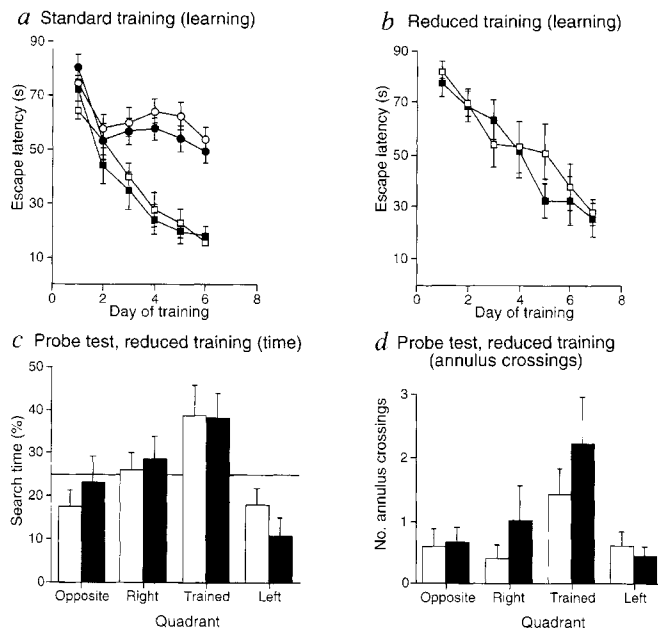


FIG. 4 Performance of *Thy-1* $+/+$ (open symbols) and $-/-$ (filled symbols) mice trained on the hidden platform version of the water maze; means $\pm$ s.e.m.. *a*, Escape latency using 4 trials per day training to find the platform hidden in a fixed position (squares, $n = 11$ for each type). Analysis of variance (ANOVA) with repeated measures showed no effect of genotype and no interaction between genotype and days of training; only the effect of days of training was significant ($F_{5,120} = 33.69$, $P < 0.001$). For a second group the platform was randomly moved on each trial (circles,

$n = 14$ for each type). For this experiment and those in *b* and *c*, identical results were obtained when the distance swum, rather than time, was plotted (not shown). *b*, Escape latency with 2 trials per day training. *Thy-1* $+/-$ ($n = 10$) and $-/-$ ($n = 9$) mice did not differ and both improved their performance during training ($F_{6,119} = 11.8$, $P < 0.01$). *c*, *d*, Performance during the probe test in the reduced training experiment, assessed either as (*c*) per cent of searching time per quadrant; the two groups did not differ and both spent more time in the trained quadrant ($F_{3,68} = 10.46$, $P < 0.001$; Newman-Keuls analysis, trained $>$ opposite and left, $P < 0.01$, trained $>$ right, $P < 0.05$); or *d*, number of annulus crossings. The two groups did not differ and both crossed the trained site more often than equivalent sites in other quadrants ($F_{3,68} = 5.13$, $P < 0.001$; Newman-Keuls analysis, trained $>$ right and left, $P < 0.01$, trained $>$ opposite, $P < 0.05$). To check whether dentate granule cells are required for spatial learning in mice, we tested a colchicine-lesioned¹³ group of mice ($n = 7$) and found their performance did not improve with training ($F < 1$). Learning was observed in a sham-lesioned group ($F_{5,65} = 7.20$, $P < 0.01$).

METHODS. Equal numbers of male and female mice, 12–15 weeks old, were tested blindly at roughly the same time each day, and were used for only one version of the test. The platform was 11.5 cm in diameter in a 150-cm diameter pool (platform/pool area/ratio 1/170) with opaque water at $24 \pm 1^\circ\text{C}$. Swimming paths were recorded with a video camera connected to a digital tracking device (Hawtrack) and processed with Water Maze software. On the first day of training, mice were placed on the platform for 20 s before testing was initiated. The starting position was varied pseudo-randomly. Each trial started with mice facing the wall of the pool and ended when they climbed onto the platform or after a maximum searching time of 90 s. At the end of each trial, they were left or placed on the platform for 15 s. In the standard training tests (fixed or random platform) animals were given 4 trials a day on 6 consecutive days within 30 min between trials. The reduced training consisted of 2 trials per day on 7 consecutive days. The probe tests were run 24 h after the last trial for 60 s starting in the opposite quadrant. Bilateral colchicine lesions were made using 12 injections of 0.12 μg colchicine in 0.9% NaCl along the septotemporal axis of the dentate gyrus, two weeks before testing.

learning (refs 11–16, and Fig. 4 legend). Our finding that *Thy-1*-deficient mice show a regionally selective suppression of LTP affords the opportunity to test whether LTP at this synaptic relay is required for spatial learning. *Thy-1* $^{-/-}$ mice learned the fixed position of a hidden platform in the watermaze⁷ at rates identical to their $+/+$ littermates. The performance of both groups was markedly superior to mice tested in the random platform version of the test where spatial cues could not be used (Fig. 4*a*). A subtle learning impairment in the *Thy-1*-mutant mice has not been hidden by overtraining, as with reduced training (Fig. 4*b*) learning was slower but identical for both groups. In a probe test at the end of training, no difference between the two groups of mice could be detected (Fig. 4*c*, *d*).

The power of reverse genetics to dissect the molecular basis for complex behaviour such as learning and memory^{17–22} is often limited by pleiotropic developmental defects²³. Such problems have been minimal here, presumably because of the late appearance of *Thy-1* in neuronal development⁵. The only defect we have seen in these mice is the absence of detectable LTP *in vivo* in the dentate gyrus, apparently the result of abnormal GABA_A-dependent inhibition in this region. Although the synaptic relay affected is the main cortical input to the hippocampus⁴, spatial learning by these mice was not compromised. Our observations therefore suggest that LTP in the entorhinal projection to the dentate gyrus may not be necessary for this form of spatial learning. □

Received 30 November 1995; accepted 4 January 1996.

- Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
- Mayford, M., Barshilai, A., Keller, F., Schacher, S. & Kandel, E. R. *Science* **256**, 638–644 (1992).
- Luthi, A., Laurent, J.-P., Figueroa, A., Müller, D. & Schachner, M. *Nature* **372**, 777–779 (1994).
- Fazeili, M. S., Breen, K., Errington, M. L. & Bliss, T. V. P. *Neurosci. Lett.* **169**, 77–80 (1994).
- Morris, R. *BioEssays* **14**, 715–722 (1992).
- Tiveron, M. C. et al. *Nature* **355**, 745–748 (1992).
- Morris, R. G. M. *Learn. Motivation* **12**, 239–260 (1981).
- Gordon, J. A., Silva, A., Morris, R., Tokugawa, Y. & Stryker, M. P. *Soc. Neurosci. Abstr.* **20**, 203.5 (1994).
- Wigstrom, H. & Gustafsson, B. *Brain Res.* **275**, 153–158 (1983).
- Sloviter, R. S. & Brisman, J. L. *J. Neurosci.* **15**, 811–820 (1995).
- Schenk, F. & Morris, R. G. M. *Exp. Brain Res.* **58**, 11–28 (1985).
- McNamara, R. K. & Skelton, R. W. *Brain Res. Rev.* **18**, 33–49 (1993).
- Whishaw, I. Q. *Behav. Brain Res.* **24**, 59–72 (1987).
- McNaughton, B. L., Barnes, C. A., Meltzer, J. & Sutherland, R. J. *Exp. Brain Res.* **76**, 485–96 (1989).
- Conrad, C. D. & Roy, E. J. *J. Neurosci.* **13**, 2582–2590 (1993).
- Cho, Y. H. & Jaffard, R. *Neurobiol. Learning Memory* **64**, 285–290 (1995).
- Silva, A. J., Stevens, C. F., Tonegawa, S. & Wang, Y. *Science* **257**, 201–206 (1992).
- Silva, A. J., Paylor, R., Wehner, J. M. & Tonegawa, S. *Science* **257**, 206–211 (1992).

- Abeliovich, A. et al. *Cell* **75**, 1263–1271 (1993).
- Bach, M. E., Hawkins, R. D., Osman, M., Kandel, E. R. & Mayford, M. *Cell* **81**, 905–915 (1995).
- Grant, S. G. N. et al. *Science* **258**, 1903–1910 (1992).
- Bourtchuladze, R. et al. *Cell* **79**, 59–68 (1994).
- Mayford, M., Abel, T. & Kandel, E. R. *Curr. Opin. Neurobiol.* **5**, 141–148 (1995).
- Swanson, L. W., Kohler, C. & Bjorklund, A. in *Handbook of Chemical Neuroanatomy*, Vol. 5 *Integrated systems of the CNS, Part I* (eds Bjorklund, A., Hokfelt, T. & Swanson, L. W.) 125–277 (Elsevier, Amsterdam, 1987).
- Forrest, D. et al. *Neuron* **13**, 325–338 (1994).
- Ransom, R. W. & Stec, N. L. *J. Neurochem.* **51**, 830–836 (1988).
- Halasy, K. & Somogyi, P. *Eur. J. Neurosci.* **5**, 411–429 (1993).
- Naquet, P. et al. *Immunol. Ser.* **45**, 99–117 (1989).
- Murphy, K. P. S. J., Williams, J. H., Bettache, N. & Bliss, T. V. P. *Neuropharmacology* **33**, 1375–1385 (1994).
- Hanse, E. & Gustafsson, B. *Eur. J. Neurosci.* **4**, 1191–1201 (1992).

ACKNOWLEDGEMENTS. This work was supported by the MRC, CNRS, The Royal Society, American Paralysis Association, International Spinal Research Trust, International Research Institute for Paraplegia. We thank H. Jani, T. Titley, D. Richards and M. Ramsey for help with some of the experiments, J. Satchell for help with the figures, and E. Hulme, S. Nonchev and P. Barber for advice.

Anterograde transport of neurotrophins and axodendritic transfer in the developing visual system

Christopher S. von Bartheld*, Margaret R. Byers†, Reg Williams‡ & Mark Bothwell*

Departments of * Physiology & Biophysics and † Anesthesiology, University of Washington, Seattle, Washington 98195-7290, USA

‡ Department of Developmental Biology, Karolinska Institute, S-17177 Stockholm, Sweden

NEUROTROPHIC factors support the differentiation and survival of neurons^{1,2} and influence properties of synaptic transmission^{3,4}. The neurotrophic hypothesis postulates a retrograde action of trophic factors: their production and release by target cells and their uptake by innervating axons⁵. Besides the retrograde route of trophic messengers, the survival of neurons and the development of synapses is thought to be also regulated by anterograde, afferent trophic signals⁶⁻¹⁰. We now show that exogenous neurotrophins are transported in the anterograde direction, from cell bodies to the axon terminals, and that the intact neurotrophin is released after anterograde transport, taken up and utilized by second-order visual neurons in the developing chick brain. These results suggest that anterogradely transported neurotrophins may play a role in synaptic plasticity and may have effects at more than one synapse beyond the initial release site.

The second-order visual neurons in the developing avian optic tectum depend on a trophic signal which is transported anterogradely by retinal ganglion cells¹¹. We examined the transport of five trophic factors in the developing retinotectal projection of

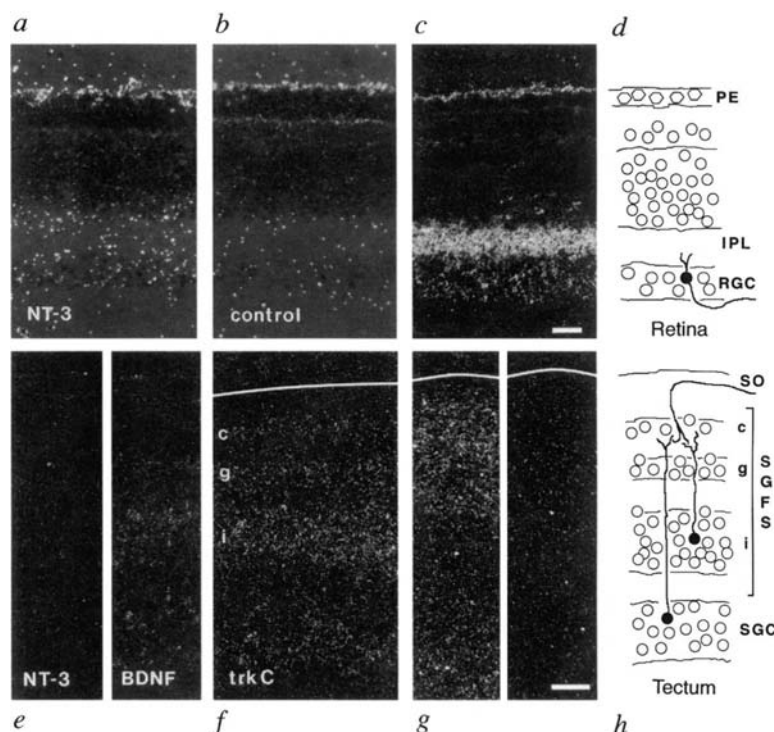
chick embryos (Figs 1, 2), including the three neurotrophins NGF (nerve growth factor), BDNF (brain-derived neurotrophic factor) and NT-3 (neurotrophin-3), as well as bFGF (basic fibroblast growth factor), and IGF-1 (insulin-like growth factor 1). All of these factors may be mediators of afferent trophic signals^{2,12-18}. The neurotrophins BDNF and NT-3 are produced in the developing retina (ref. 19 and Fig. 1a), and trkB and trkC, the tyrosine kinase receptors for BDNF and NT-3 (refs 20, 21), are expressed in the optic tectum (trkC data shown in Fig. 1f). In addition, the neurotrophin receptor, p75, which binds all neurotrophins with similar affinity, is expressed in the optic tectum²². NT-3 binds with high affinity in the tectal neuropil rather than in layers containing neuronal cell bodies (compare Fig. 1g and h), suggesting that anterogradely transported NT-3 may bind to trkC receptors on optic nerve fibres or on the dendrites of second-order visual neurons.

The eyes of 15–16-day-old chick embryos were injected with 10–100 ng radio-iodinated NGF, BDNF, NT-3, IGF-1 or bFGF. The axonal transport by retinal ganglion cells (RGC) to their target, the optic tectum, was quantified 2–64 h after injection. Maximal transport occurred by 20 h (Fig. 2a–c). NT-3 was transported consistently and at a significantly higher level than the other trophic factors examined (Fig. 2c). About 10–20% of the NT-3 evident in RGC bodies was transported anterogradely to the tectum, based on measurements of radioactivity in the two sites and the analysis of autoradiographic profiles (Fig. 2f). On average, each RGC transported about 1,000 radioactive NT-3 molecules to the tectum. Most NT-3 accumulated in the stratum opticum (SO) and the stratum griseum fibrosum et superficiale (SGFS), layers a–h (Fig. 2b).

Transport of NT-3 appeared to be receptor-mediated as it was competed by excess cold NT-3 (Fig. 2c) and denatured NT-3 was not transported (Fig. 2d). Anterograde transport of NT-3 was significantly more efficient than that of BDNF. This difference is not likely to be due to differences in the diffusion of the two neurotrophins in the retina, because they accumulated in a

FIG. 1 Sections through the retina (a–d) and the optic tectum (e–h) of 16-day-old chick embryos. a, *In situ* hybridization shows that the neurotrophin NT-3 is expressed in the retina. b, Adjacent section hybridized with a sense (control) probe. c, Radio-iodinated NT-3 accumulates in the inner plexiform layer (IPL) and retinal ganglion cell layer (RGC) after intraocular injection. d, The laminar organization of the retina is indicated. PE, pigment epithelium. e, Lack of NT-3 mRNA expression in the tectum (left). The adjacent section hybridized with a probe for BDNF shows label in layers SGFS g, i and SGC (right). f, *In situ* hybridization shows that the NT-3 receptor, trkC, is expressed in layers SGFS c, g, i and SGC of the optic tectum. Similar data were obtained for trkB (data not shown). g, Binding of radio-iodinated NT-3 in the presence of excess cold NGF (left) or excess cold NT-3 (right). Iodinated NT-3 binds with high affinity predominantly in the neuropil region (compare with panel h), suggesting that dendrites and axons in the tectum rather than cell bodies of second-order visual neurons contain the trkC receptor. Similar data were obtained with iodinated BDNF (data not shown). h, The laminar organization of the tectum is indicated. SO, stratum opticum; SGFS, stratum griseum et fibrosum superficiale with sub-layers c, g and i; SGC, stratum griseum centrale. Scale bars: a–c, 20 µm; e–g, 200 µm.

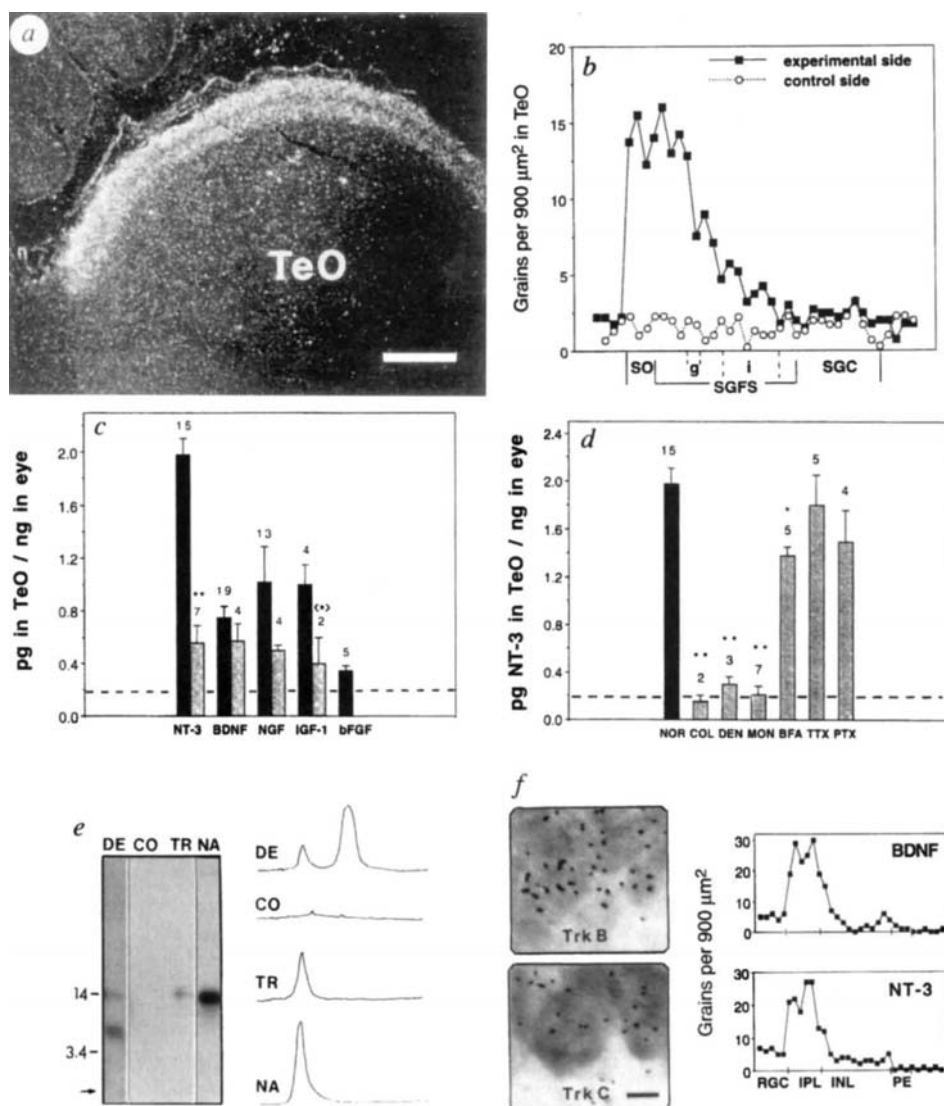
METHODS. A 315-bp fragment complementary to chicken NT-3 DNA and a 527-bp fragment complementary to chicken BDNF DNA were used to generate antisense riboprobes for *in situ* hybridization (chicken NT-3 and BDNF DNA: courtesy of P. Maisonpierre and G. Yancopoulos). *In situ* hybridization was performed with ³⁵S-labelled riboprobes on paraffin sections as described²². Frozen sections were hybridized with a probe that recognizes mRNA encoding the kinase-containing form of trkC²⁸. NT-3 was iodinated with lactoperoxidase²³. Binding studies were performed with 2 ng ml⁻¹ NT-3 at room temperature in the presence or



absence of 1,000-fold excess cold NT-3 or NGF as described for NGF²⁹. The pattern of binding in the presence of excess cold NGF was virtually identical to that with iodinated NT-3 only (data not shown).

FIG. 2a, Anterograde transport of iodinated NT-3 to the optic tectum (TeO) and distribution in superficial tectal layers, 20 h after injection into the eye of a 15-day-old chick embryo. Autoradiography of emulsion-coated section through the tectum. Label distributes throughout the zone of termination of retinotectal axons. Scale bar, 1 mm. *b*, Grain profile of anterogradely transported iodinated NT-3 in the tectum. The number of autoradiographic silver grains was plotted over the tectal layers SO, SGFS and SGC. *c*, The amount of transport was quantified for neurotrophin-3 (NT-3), brain-derived neurotrophic factor (BDNF), nerve growth factor (NGF), insulin-like growth factor 1 (IGF-1) and basic fibroblast growth factor (bFGF). Transport of NT-3 to the tectum (TeO) is significantly higher than that of any of the other trophic factors examined ($P < 0.001$, ANOVA, multiple comparisons). The dashed line indicates background due to leakage into systemic circulation. Grey bars show the amount of transport when 30–150-fold excess cold homologous factor was co-injected into the eye. Transport of NT-3 is largely competed out by excess cold NT-3 ($P < 0.005$, *t*-test) and thus appears to be receptor-mediated. Transport of NGF is not competed significantly by excess cold NGF ($P < 0.375$); competition of IGF-1 transport is significant at $P < 0.05$. The uptake of labelled neurotrophins in RGC cell bodies was reduced by 30–50% with excess cold neurotrophin (data not shown). Bars represent s.e.m.; numbers over bars are the number of independent experiments (in *c*, *d*). *d*, Transport of NT-3 was further investigated. Transport of intact NT-3 is completely abolished by co-injection of colchicine (COL, 7 μ g) or monensin (MON, 9 μ g). Denatured NT-3 (DEN) is not transported. Transport is only marginally affected by brefeldin A (BFA, 15 μ g), tetrodotoxin (TTX, 0.3 μ g), or pertussis toxin (PTX, 0.2 μ g). NOR, normal condition. *e*, Virtually all NT-3 arrives in the tectum intact, as shown by SDS–polyacrylamide gel electrophoresis and laser densitometry of the lower half of the gels. DE, sample from tectal homogenate to which trypsin-degraded 125 I-NT-3 was added (positive control showing that degraded NT-3 would have been detected); CO, sample from control side (ipsilateral tectum); TR, sample from the experimental tectum (containing transported NT-3); NA, native NT-3; arrow indicates the dye front. *f*, Photomicrographs showing the higher level of expression of trkB than trkC in retinal ganglion cells (RGC) of a 16-day-old chick embryo. Scale bar, 5 μ m. BDNF, the preferred ligand of trkB, and NT-3, the preferred ligand of trkC, accumulated at similar levels in RGC neurons, as shown by grain-density profiles of sections through the retina after injection of the radio-labelled neurotrophins in the eye.

METHODS. Trophic factors were iodinated with lactoperoxidase as described²³. Specific activities were 73.9–129.5 c.p.m. per pg (NT-3), 83.6–175 c.p.m. per pg (BDNF), 61–112 c.p.m. per pg (NGF), 128.7 c.p.m. per pg (IGF-1) and 102.7 c.p.m. per pg (bFGF). Chicken embryos of 15 days of incubation received intraocular injections of 10–100 ng iodinated trophic factors. Embryos were anaesthetized 20 h after injection and perfused transcardially with 4% paraformaldehyde. Disintegrations per min were documented in a γ -radiation counter after dissection of the injected eye and the midbrain. The accumulation of trophic factors in the RGC cell bodies was estimated from the c.p.m. in the entire eye; the eye was sectioned for autoradiography, and the distribution of silver grains plotted over the vitreous and the different layers of the retina. The percentages of radioactivity in different structures of the eye were calcu-



lated from these profiles. To quantify the transport to the optic tectum, the amount of radioactivity due to retrograde transport to the isthmo-optic nucleus was subtracted from total counts in the midbrain. The amount of retrograde transport to the isthmo-optic nucleus²³ was measured by correlating grain density and disintegration counts in midbrains from animals co-injected with monensin in the eye. Monensin abolishes anterograde transport (*d*) without affecting retrograde transport (data not shown). The amount of trophic factor in the tectum is expressed as pg of anterogradely transported trophic factor per ng of injected factor present in the eye at the time of death (*c*, *d*). Each value represents the average of 2–19 independent experiments (*n*, indicated). Error bars represent s.e.m. Significance levels (**, $P < 0.005$; *, $P < 0.01$; (*), $P < 0.05$) were determined by *t*-test. For SDS gel analysis, tecta labelled by anterograde transport were homogenized in triple detergent lysis buffer (containing 0.1% SDS, 1% Nonidet P-40 and 0.5% sodium deoxycholate), centrifuged, and samples of the supernatant (containing most of the radioactivity) were diluted in gel-loading buffer containing SDS, boiled for 4 min, and loaded onto a 15% SDS–polyacrylamide gel, run for 4 h, stained with Coomassie blue, dried and exposed for 2–12 weeks using intensifier screens. In control experiments, free Na^{125}I , trypsin-digested ^{125}I -NT-3 or intact ^{125}I -NT-3 was added to the homogenate. These control experiments showed that degraded ^{125}I -NT-3 would readily be detected on the gel; free Na^{125}I would not be detected. Therefore, control experiments were performed in which tecta containing transported ^{125}I -NT-3 were homogenized, or free Na^{125}I , or intact ^{125}I -NT-3 were added to (unlabelled) tectal homogenates. Most of the radioactivity transported to the tectum was precipitated by trichloroacetic acid (TCA) (incorporated into protein). As degraded protein is visible in the control SDS gel (*e*), we conclude that most of the radioactivity transported to the tecta represents intact NT-3.

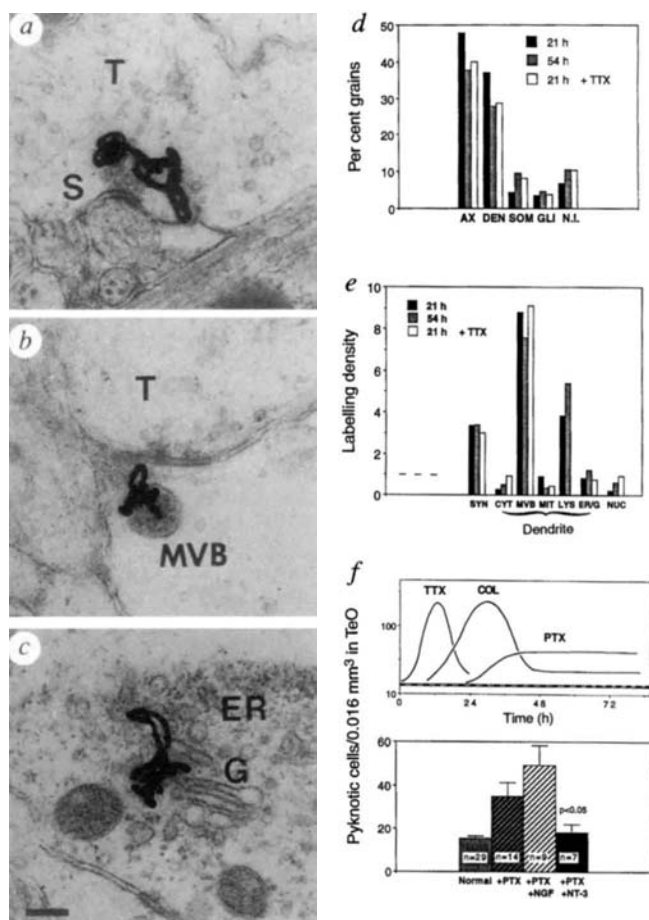
similar fashion in the IPL and RGC layers (Figs 1c, 2f), and there was no difference between BDNF and NT-3 in the retrograde transport to the isthmo-optic nucleus²³. Anterograde transport of exogenous neurotrophins was only slightly reduced by the drug brefeldin A (BFA, Fig. 2d) which prevents the transfer of newly synthesized proteins from the endoplasmic reticulum to the Golgi complex²⁴. Transport was not affected by the sodium-channel blocker tetrodotoxin (TTX, Fig. 2d), which abolishes electrical activity¹¹. Treatment with the Golgi-disrupting drug monensin²⁵ prevented transport entirely (Fig. 2d). The effect of monensin is consistent with the hypothesis that exogenous NT-3 joins the pathway of newly synthesized proteins in the Golgi complex of RGC. More than 95% of exogenous NT-3 that arrived in the tectum was intact, as revealed by laser densitometry of SDS-polyacrylamide gels (Fig. 2e).

To test whether neurotrophins are released after anterograde transport or remain within RGC axons, we examined the distribution of NT-3 in the tectum by quantitative electron microscopic autoradiography. Radioactive NT-3 accumulated on the presynaptic side of synaptic contacts (Fig. 3a). About 40% of the transported NT-3 was transferred to dendrites and cell bodies of second-order visual neurons (Table 1 and Fig. 3b-d). The distribution of NT-3 was similar at 21 and 54 h after injection and co-injection of tetrodotoxin (TTX) had no significant effect, suggesting that the transport and axodendritic transfer of NT-3 is independent of electrical activity (Fig. 3d, e and Table 1). Within dendrites, nearly half of the NT-3 was associated with

multivesicular bodies (Fig. 3b). Another major fraction labelled the endoplasmic reticulum/Golgi system (Fig. 3c). Multivesicular bodies (in distal and proximal dendrites) and lysosomes (in proximal dendrites) had the highest labelling density (Fig. 3e). The fact that internalized NT-3 initially accumulates in multivesicular bodies is consistent with the notion that these organelles are involved in sorting of internalized trophic molecules²⁶. NT-3 may be sorted at postsynaptic sites or during retrograde dendritic transport into either of two pathways: degradation or re-use (anterograde transport to the terminals).

Anterogradely transported NT-3 distributes after axodendritic transfer in a way that is very similar to that of NGF after retrograde transport¹⁵. This suggests that anterogradely transported NT-3 may be used for trophic signalling. To test this idea directly, we made use of the fact that some neurotoxins can interrupt the trophic signalling of presumptive endogenous trophic factors without affecting anterograde transport of exogenously applied neurotrophins. Neurons in the stratum griseum centrale (SGC) of the optic tectum require trophic signals from the retina¹¹. Interruption of transport by colchicine or abolition of electrical activity increases cell death (Fig. 3f), as does the intraocular injection of pertussis toxin (PTX). PTX presumably interferes with the expression or secretion of trophic factors in the retina²⁷. Internalization and anterograde transport of exogenous NT-3 was not significantly affected by PTX (Fig. 2d). Consistent with our hypothesis, PTX-treated tectal neurons were rescued by intraocular, exogenous NT-3, but not by NGF (Fig. 3f). The simplest

FIG. 3 Electron microscopic autoradiography of anterogradely transported NT-3 in the optic tectum (a-e) and test of functional effects of NT-3 in the tectum (f). a, b, Anterogradely transported NT-3 accumulates in synapses (S) and in multivesicular bodies (MVB) of dendrites in tectal neurons. T, terminal. Layer d of the stratum griseum et fibrosum superficiale (SGFSd). The MVB in b is in close vicinity to a synapse in layer SGFSd of the optic tectum. c, Silver grains associated with the endoplasmic reticulum (ER) and Golgi system (G) in a proximal dendrite in layer SGFSd. Scale bar (a-c), 200 nm. d, About 40% of silver grains in the tectum are over dendrites (DEN) and the neuronal soma (SOM) of second-order visual neurons. This percentage is similar at 21 and 54 h, and is not affected by co-injection of TTX. AX, axons; GLI, glial cells; NI, not identified. e, Synapses (SYN), multivesicular bodies (MVB), and lysosomes (LYS) have significantly higher labelling densities than the cytosol (CYT), mitochondria (MIT), the endoplasmic reticulum/Golgi system (ER/G) and the nuclei of neurons (NUC). Rare silver grains on the control side are randomly distributed (data not shown). The labelling density was calculated as the percentage of total grains divided by the percentage of total area of each organelle. Random distribution has a labelling density of 1.0 (dashed line). f, Intraocular injection of tetrodotoxin (TTX, 0.3 µg)¹¹, colchicine (COL, 7 µg)¹¹, or pertussis toxin (PTX, 0.2 µg) increases cell death in the SGC of the experimental optic tectum (upper panel), based on assessment of cell death between 5 and 88 h after injection (TTX: 5, 9, 12, 13, 17, 20, 24, 30 h—compare with ref. 11; COL: 9, 13, 17, 20, 31, 48, 61, 88 h—compare with ref. 11; PTX: 13, 20, 24, 39, 48, 62, 78 h). Bold base line represents normal developmental cell death. Pyknotic cells in the SGC were identified according to established criteria¹¹. The cell death induced by PTX 48 h after injection was largely reduced by co-injection of 300 ng NT-3 in the same eye, but not by co-injection of 300 ng NGF (lower panel). Bars represent S.E.M. The P value for +PTX/+NT-3 (against +PTX) was determined by t-test and Mann-Whitney test. To avoid examiner bias, cells were counted blind as to treatment. The number of independent experiments is indicated. METHODS. Three chick embryos were injected with 100 ng iodinated NT-3 into the vitreous body of one eye, killed after 21 or 54 h, and several samples of the experimental and control tectum were processed for electron microscopic autoradiography using calibrated monolayer emulsion films (Ilford L4) as described³⁰. Each sample from the experimental side was examined in 4 systematic traverses, from the stratum opticum (SO) to the stratum griseum centrale (SGC), and each silver grain was photographed using a Philips 410 TEM. A total of 530 silver grains were photographed and the underlying structure(s) identified. A probability circle analysis of the source was performed (Table 1) and the fractional areas and the labelling densities of organelles were determined as described¹⁵. Rare (background) silver grains on the control side (Fig. 2b) were analysed in the same way. In e and in Table 1, the category 'synapse' includes pre- and postsynaptic label



(within 300 nm from the synaptic cleft). About 65% of the synaptic label was presynaptic; 35% was postsynaptic. 'Cytosol' was scored when no organelle was obvious within a radius of 300 nm within dendrites. The 'glia' category combines astroglia and oligodendroglia.

TABLE 1 Labelled structures in the optic tectum after transport of ^{125}I -labelled NT-3

Structure	Fractional area	Per cent grains			Labelling density		
		21 h	+TTX	54 h	21 h	+TTX	54 h
Axon	25.2	28.0	27.8	31.5	1.11	1.10	1.25
Terminal axon	9.7	10.8	12.1	6.2	1.11	1.25	0.64
Synapse	2.7	9.0	8.1	9.1	3.33	3.00	3.37
Distal dendrite	27.2	28.2	22.0	21.6	1.04	0.81	0.79
Proximal dendrite	10.3	8.8	6.6	6.2	0.85	0.64	0.60
Neuron soma	9.3	4.3	8.1	9.5	0.46	0.87	1.02
Astroglia	6.8	3.5	3.7	4.6	0.51	0.54	0.68
Unidentified	7.7	6.7	10.3	10.4	0.87	1.34	1.35
Number of grains sampled		250	130	150	—	—	—

interpretation of these data is that NT-3 is an anterogradely transported trophic messenger, although we cannot rule out an alternative interpretation, that NT-3 acts trophically on RGCs to enhance the anterograde transport and release of a second trophic factor. TTX-induced cell death was not prevented by NT-3 injected in the eye 0 or 18 h before TTX (data not shown). This suggests that TTX induces cell death by a different mechanism from PTX, and that the response to NT-3 may require activation.

Our study on the anterograde transport of exogenously introduced neurotrophins, together with recent data on the anterograde transport of endogenous neurotrophins³¹, provides direct evidence that neurotrophins may act as anterograde trophic messengers. Enhanced survival may be but one of several functions of anterogradely transported neurotrophins. BDNF and NT-3 modify synaptic transmission of developing synapses^{3,4}, and these neurotrophins may have important functions in the stabilization of appropriate synapses and the loss of inappropriate ones during development^{5,7,8}. The modification of synaptic structure and efficacy may also play a role in synaptic plasticity of the adult nervous system, and particularly in processes related to learning and memory². □

Received 7 August; accepted 6 December 1995.

- Barde, Y. A. *Neuron* **2**, 1525–1534 (1989).
- Thoenen, H. *Trends Neurosci.* **14**, 165–170 (1991).
- Lohof, A. M., Ip, N. Y. & Poo, M. M. *Nature* **363**, 350–353 (1993).
- Kang, H. & Schuman, E. M. *Science* **267**, 1658–1662 (1995).
- Purves, D. *Body and Brain. A Trophic Theory of Neural Connections* (Harvard University Press, Cambridge, 1988).
- Cowan, W. N. in *Contemporary Research Methods in Neuroanatomy* (eds Nauta, W. J. H. & Ebesson, S. O. E.) 217–251 (Springer, New York, 1970).
- Changeaux, J. P. & Danchin, A. *Nature* **264**, 705–712 (1976).
- Vaughn, J. E. *Synapse* **3**, 255–285 (1989).
- Clarke, P. G. H. *Concepts Neurosci.* **2**, 201–219 (1991).
- Oppenheim, R. W. A. *Rev. Neurosci.* **14**, 453–501 (1991).
- Catsicas, M., Pequignot, Y. & Clarke, P. G. H. *J. Neurosci.* **12**, 4642–4650 (1992).
- Korsching, S. J. *Neurosci.* **13**, 2739–2748 (1993).
- Ferguson, I. A., Schweitzer, J. B. & Johnson, E. M. Jr *J. Neurosci.* **10**, 2176–2189 (1990).
- Nieto-Bona, M. P., Garcia-Segura, L. M. & Torre-Aleman, I. J. *Neurosci. Res.* **36**, 520–527 (1993).
- Claude, P., Hawrot, E., Dunis, D. A. & Campenot, R. B. *J. Neurosci.* **2**, 431–442 (1982).
- Wayne, D. B. & Heaton, M. B. *Dev. Biol.* **138**, 484–498 (1990).
- Ernfors, P., Wetmore, C., Olson, L. & Persson, H. *Neuron* **5**, 511–526 (1990).
- Schecterson, L. C. & Bothwell, M. *Neuron* **9**, 449–463 (1992).
- Herzog, K. H., Bailey, K. & Barde, Y. A. *Development* **120**, 1643–1649 (1994).
- Meakin, S. & Shooter, E. M. *Trends Neurosci.* **15**, 323–331 (1992).
- Bothwell, M. A. *Rev. Neurosci.* **18**, 223–253 (1995).
- von Bartheld, C. S., Heuer, J. G. & Bothwell, M. J. *comp. Neurol.* **310**, 103–129 (1991).
- von Bartheld, C. S. et al. *Neuron* **12**, 639–654 (1994).
- Smith, R. S., Hammerslag, R., Snyder, R. E., Chan, H. & Bobinsky, J. J. *Neurochem.* **62**, 1698–1706 (1994).
- Hammerslag, R. & Stone, G. C. *Trends Neurosci.* **5**, 12–15 (1982).
- Felder, S. et al. *Cell* **61**, 623–634 (1990).
- Konrad, R. J. et al. *J. Biol. Chem.* **270**, 12869–12876 (1995).
- Williams, R., Bäckström, A., Kullander, K., Hallböök, F. & Ebendal, T. *Eur. J. Neurosci.* **7**, 116–128 (1995).
- Richardson, P. M., Verge, V. M. K. & Riopelle, R. J. *J. Neurosci.* **6**, 2312–2321 (1986).
- Byers, M. R. *Brain Res.* **75**, 97–113 (1974).
- Zhou, X. F. & Rush, R. A. *Neuroscience* (in the press).

ACKNOWLEDGEMENTS. BDNF and NT-3 were kindly provided by Regeneron, bFGF by Chiron Corporation, and chicken BDNF and NT-3 DNA by P. Maisonpierre and G. Yancopoulos (Regeneron). We thank E. Gunther, R. Rush and X.-F. Zhou for sharing unpublished data; L. Schecterson for subcloning; S. Anderson and D. Cunningham for technical assistance; and Y. Barde, P. Clarke, E. Johnson, W. Mobley, R. Oppenheim, F. Peale, C. Shatz, H. Thoenen and L. Westrum for their comments. Our work was supported by grants from the NIH.

The inward rectification mechanism of the HERG cardiac potassium channel

Paula L. Smith, Thomas Baukrowitz & Gary Yellen*

Department of Neurobiology, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114-2698, USA

A HUMAN genetic defect associated with 'long Q-T syndrome', an abnormality of cardiac rhythm involving the repolarization of the action potential, was recently found to lie in the HERG gene, which codes for a potassium channel¹. The HERG K⁺ channel is unusual in that it seems to have the architectural plan of the depolarization-activated K⁺ channel family (six putative transmembrane segments), yet it exhibits rectification like that of the inward-rectifying K⁺ channels, a family with different molecular structure (two transmembrane segments)^{2–4}. We have studied HERG channels expressed in mammalian cells and find that this inward rectification arises from a rapid and voltage-dependent inactivation process that reduces conductance at positive voltages. The inactivation gating mechanism resembles that of C-type inactivation, often considered to be the 'slow inactivation' mechanism of other K⁺ channels. The characteristics of this gating suggest a specific role for this channel in the normal suppression of arrhythmias.

HERG K⁺ channels expressed in mammalian cells have the same basic properties previously reported for channels expressed in oocytes^{3,4}: they required depolarization for their activation, but they carried a relatively small outward current (Fig. 1a). On repolarization to a negative membrane potential, the conductance increased to a large value within a few milliseconds, before returning slowly to the resting closed state. The peak currents measured by stepping to various voltages after activation showed inward rectification (that is, the inward currents were substantially larger than the outward) (Fig. 1b). This inward rectification has been proposed to result from an 'inactivation' mechanism that turns channels off at positive potentials but quickly recovers at negative potentials^{3–6}, producing a rising phase in the inward current. Unlike the clear time-dependent inactivation of other voltage-activated K⁺ channels like *Shaker*, this inactivation was thought to be kinetically invisible for depolarizing steps because it is much faster than activation.

The onset of this inactivation process was revealed by a modified voltage-clamp pulse protocol. After a long depolarizing pulse to activate the channels, the membrane was hyperpolarized

* To whom correspondence should be addressed at the Massachusetts General Hospital.

(Fig. 1c). Channels recovered rapidly from inactivation and then began to close slowly, but before they closed appreciably the membrane was again depolarized. At this time the channels were still activated (because they had not yet closed), but the inactivation process had to re-occur. This 're-inactivation' was seen as a single exponential decline in current, returning to the same steady-state level seen in the first depolarizing pulse.

Without this inactivation, the channels did not exhibit inward rectification. Channels were activated by depolarization and then hyperpolarized to -80 mV to allow them to recover from inactivation. From this starting point, an instantaneous current-voltage (I - V) relationship obtained by steps to varying voltages was roughly linear (Fig. 2a, b). This method measured the current through channels before inactivation gating was allowed to adjust to the new voltage. Thus, inactivation gating produces inward rectification in a channel with non-rectifying permeation properties.

We measured the steady-state voltage dependence of inactivation to see if it matched the voltage-dependent rectification. After a long depolarizing pulse to activate channels, the membrane voltage was stepped briefly to various test voltages and then to a constant measurement potential of $+20$ mV (Fig. 2c). During the brief step, the inactivation process relaxed rapidly to the steady-state level appropriate to the test potential. The initial current on stepping to $+20$ mV then gave the relative number of open channels, which was large at negative voltages that relieved inactivation, but smaller at more positive voltages where steady-state inactivation was significant. For very negative voltages, we corrected for the fast closing that occurred during the brief hyperpolarizing step (see Fig. 2 legend). The data could be fitted by a Boltzmann function with half inactivation at about -90 mV and an effective valence of about 1 (Fig. 2d).

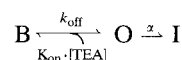
If inward rectification results from this inactivation process, then the rectifying I - V should be given by the product of the open channel I - V (Fig. 2b) and the steady-state inactivation function (Fig. 2d). To test this, we re-plotted the peak I - V data of Fig. 1a, b with corrections at negative voltages for the de-activation process, and compared the rectification with the prediction (Fig. 2e). The match between the two shows that the inward rectification of HERG channels can be quantitatively accounted for by the properties of the inactivation process.

We then investigated the molecular basis of the HERG inactivation process. One mechanism for inward rectification used by some members of the inwardly rectifying K^+ channel family, is a voltage-dependent channel blockade by intracellular Mg^{2+} (refs 7-10). If the onset of this blockade were sufficiently slow, it could appear as a time-dependent inactivation process. But we found that Mg^{2+} did not play a role in HERG inactivation. The presence or absence of Mg^{2+} in the intracellular solution perfusing an excised inside-out patch made no difference in the rate or extent of inactivation (Fig. 3a).

Inactivation of another voltage-activated K^+ channel, the *Shaker* K^+ channel from *Drosophila*, occurs by two distinct molecular mechanisms: N-type inactivation, which occurs by an intracellular ball-and-chain mechanism¹¹⁻¹³, and C-type inactivation, which involves a conformational change at the outer mouth of the channel¹⁴⁻¹⁶. The classical K^+ channel blocker tetraethylammonium (TEA) can be used to distinguish the two mechanisms: blockade by intracellular TEA inhibits N-type inactivation, whereas blockade by extracellular TEA inhibits C-type inactivation¹⁵.

We tested the effects of TEA on the HERG inactivation process. Intracellular TEA applied to an inside-out patch reduced the HERG current with an apparent affinity of about 0.2 mM, but there was no change in the time course of inactivation (Fig. 3b). This effectively ruled out an N-type inactivation mechanism, and indeed any mechanism that relies on an intracellular pore blocker. Any such mechanism should compete with the binding of TEA to the intracellular mouth, and it would thus be slower in the presence of TEA.

By contrast, extracellular TEA did interfere with inactivation. Extracellular application of 100 mM TEA reduced the peak current and slowed the time course of inactivation (Fig. 3c). The simplest explanation for this combination of effects, which is also seen for the C-type inactivation mechanism of *Shaker* channels¹⁵, is that blocked channels (B) cannot inactivate (I) (O, open):



If binding and dissociation of the blocker are rapid (at equilibrium) relative to the rate of inactivation (α), then the apparent rate

FIG. 1 Inward rectification of HERG and the HERG inactivation process. a, HERG currents transiently expressed in an HEK293 cell were activated by a long depolarizing pulse to $+20$ mV, and then stepped to various voltages to show inward rectification. b, Peak I - V from the currents in a. c, Method used to show the onset of HERG inactivation. After a long step to $+20$ mV to activate the channels, repolarization to -80 mV produced rapid recovery from inactivation. If the voltage was maintained at -80 mV, channels proceeded to close by de-activation. If the voltage was stepped back to $+20$ mV after 30 ms, the recovered channels produced a large current that then declined through the inactivation process ($\tau \sim 11$ ms). Currents in c were measured from an excised outside-out patch.

METHODS. The HERG complementary DNA was subcloned into the GW1-CMV vector under control of the human cytomegalovirus promoter¹⁵. This cDNA contained two point mutations (V198E and P202L) compared with the published HERG sequence (G. Robertson, personal communication). Transfection and electrophysiology methods were as previously described^{22,24}. The basic internal solution was 160 mM KCl, 1 mM EGTA, 0.5 mM $MgCl_2$, 10 mM HEPES, pH 7.4 and in the case of the zero internal Mg^{2+} experiments, $MgCl_2$ was replaced by 5 mM EDTA. The external solution contained 150 mM NaCl, 10 mM KCl, 3 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM HEPES, pH 7.4. The HEK293 cells express variable amounts of an endogenous outward rectifier K^+ channel, which is blocked by external TEA with an affinity of about 1 mM (data not shown). We did not use cells or patches where the endogenous current was substantial compared with HERG expression.

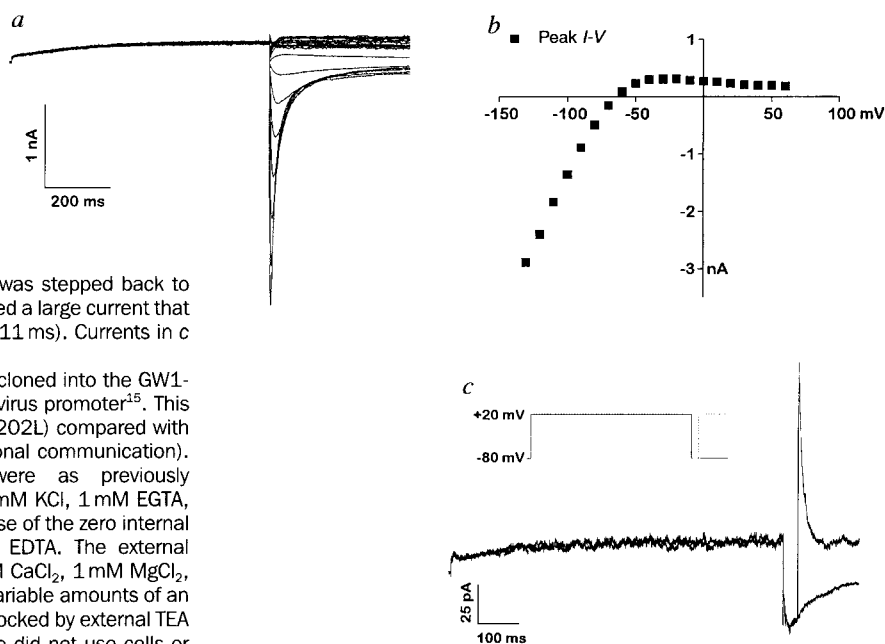


FIG. 2 The inactivation process accounted for inward rectification. *a*, An instantaneous $I-V$ protocol showed that the open-channel $I-V$, before the onset of inactivation, was linear. *b*, Plot of the instantaneous $I-V$ measured at the time of the arrow in *a*. *c*, Method used to assess steady-state inactivation. After allowing inactivation to relax to steady state at various voltages, the membrane voltage was stepped to +20 mV to assess the relative number of channels available to activate. *d*, Steady-state inactivation from *c*. The initial current (arrow in *c*) was plotted as a function of the previous voltage step ($\circ$). At negative voltages, the currents declined because significant closing of channels occurred through de-activation. This was corrected for by extrapolating the exponential falling phase (measured in full at the end of the procedure, but not shown; see Fig. 1a) back to the start of the negative voltage step, and applying the same relative correction to the initial outward current. The corrected inactivation curve ($\blacksquare$) was well-fitted by a Boltzmann function (solid line) with a voltage midpoint of -90 ± 6 mV and an effective valence of -1.0 ± 0.1 . *e*, Comparison between the measured inward rectification and the rectification predicted from the inactivation curve. Inward rectification ($\blacksquare$) was measured as in Fig. 1b and then corrected for closing at negative voltages (as in *d*). The prediction was made by multiplying the instantaneous $I-V$ (from *b*) by the inactivation function (from *d*). Values from three experiments were normalized to equal one at -100 mV and then averaged; the error bars indicate the s.e.m.

METHODS. Back-extrapolation of the closing phase may not be precisely correct if inactivated channels do not close at the same rate. Nevertheless we think that the error was small as fits constrained only by the uncorrected data in the voltage range -80 to $+40$ gave roughly the same parameters.

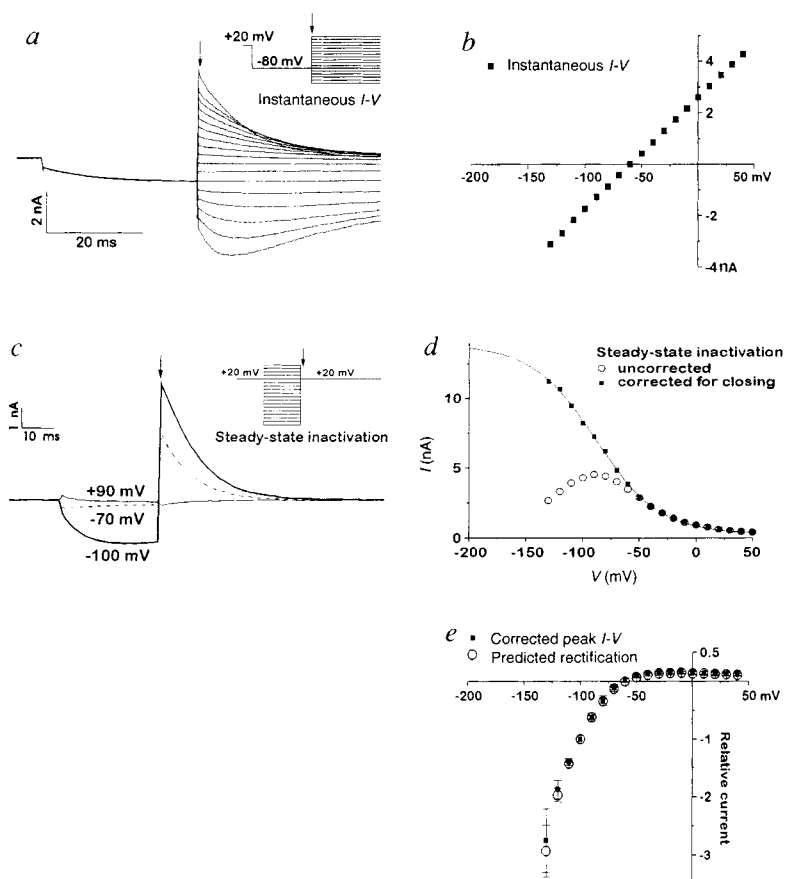
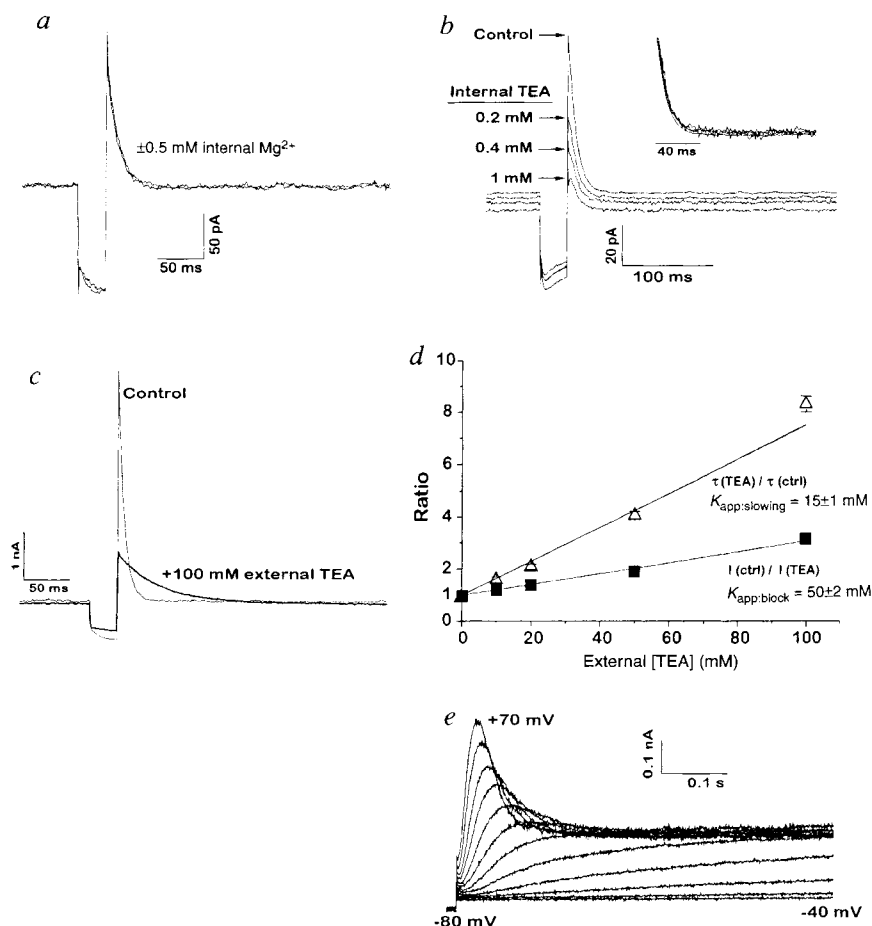


FIG. 3 Effects of Mg^{2+} ions and TEA on the HERG inactivation process. *a*, HERG currents from an inside-out patch perfused with 0.5 mM Mg^{2+} or with 5 mM EDTA. The initial voltage was +20 mV, followed by a 30-ms pulse to -80 mV, with re-inactivation at +20 mV. *b*, HERG currents from an inside-out patch perfused with various concentrations of internal TEA. Voltage steps as in *a*. The horizontal arrows indicate the initial peaks, which declined as internal TEA was raised. All currents inactivated with a time constant of about 9 ms. If the currents were normalized at the peak, they superimposed (inset). *c*, External TEA blocked the peak current and slowed inactivation. Voltage steps as in *a*. *d*, Effect of external TEA on the time constant of inactivation (Δ) and on the peak initial current at +20 mV ($\blacksquare$). In each case, the fold-change was plotted against [TEA]. K_{app} for TEA in each case was given by the reciprocal of the slope. Mean and range of two experiments. *e*, Activating voltage family in the presence of 100 mM external TEA. At positive voltages, activation was faster than the inactivation (which had been slowed by the TEA). METHODS. The intrinsic voltage dependence of TEA blockage was small or nonexistent (data not shown). If it were large, the relationship in the text would not be correct, because the redistribution between gating states would occur at -80 mV but the blockade would be measured finally at +20 mV.



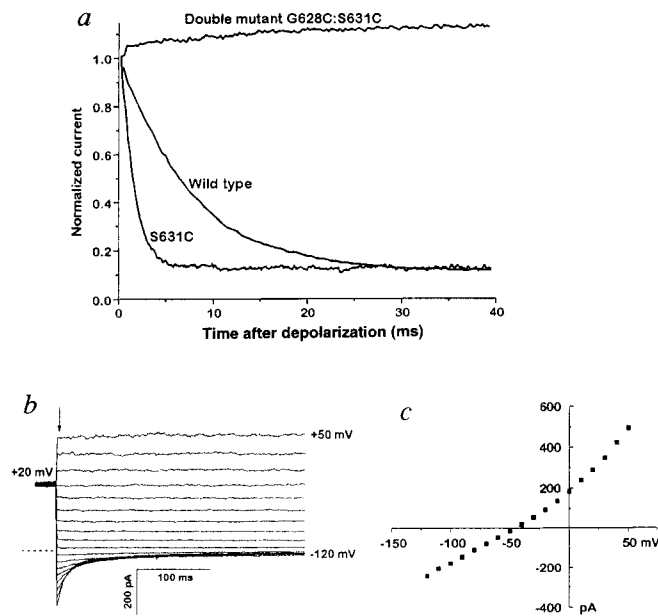


FIG. 4 Effects of two pore mutations on HERG inactivation. *a*, Comparison of inactivation for wild-type HERG channels and the mutant channels S631C and G628C:S631C. Inactivation was observed at +20 mV using methods like those in Fig. 3. For wild type, the preceding step to -80 mV lasted 35 ms; for the mutants, it was 15 ms. *b*, Absence of inward rectification in the HERG double mutant G628C:S631C, using the method of Fig. 1*b*. Dotted line indicates zero current level. *c*, *I*-*V* relationship for the double mutant, measured at the time indicated by the arrow in *b*.

of inactivation will be

$$\alpha_{app} = \alpha \cdot \left[\frac{O}{O + B} \right] = \alpha \cdot \left[\frac{1}{1 + [TEA]/K_d} \right]$$

where $K_d = k_{off}/k_{on}$ and the time constant for inactivation in TEA (assuming that recovery is much slower than onset) is $\tau_{TEA} = \tau_{ctrl} \cdot (1 + [TEA]/K_d)$, where $\tau_{ctrl} = \alpha^{-1}$. Plotting τ_{TEA}/τ_{ctrl} as a function of [TEA] gave this straight-line relationship (Fig. 3*d*), with a K_d value of 15 ± 1 mM.

We compared this with the affinity estimated from current inhibition. A simple plot of current reduction as a function of [TEA] gave a much lower apparent affinity of 50 ± 2 mM (Fig. 3*d*), but this measurement was tainted by the interaction between TEA and inactivation. In the presence of TEA, channels equilibrate during the prepulse to -80 mV among three states—open, blocked and inactivated—and the redistribution of channels from the inactivated state reduces the apparent blockade. The apparent affinity measured under these conditions should be $K_{app} = K_d \cdot (1 + K_I)$, where K_d is the true dissociation constant and K_I is the equilibrium constant for inactivation at -80 mV. From the steady-state inactivation measurement of Fig. 2*c*, $K_I = 1.7 \pm 0.5$, and the true K_d estimated from the current inhibition was 19 ± 4 mM, in reasonable agreement with the K_d for the slowing of inactivation.

Because TEA slows the effective rate of HERG inactivation, it can be used as an alternative tool for separating the two gating processes. In the presence of TEA, and particularly at very positive voltages, activation became faster than inactivation and the outward HERG current resembled that of a 'normal' voltage-activated channel like *Shaker*: rapid voltage-dependent activation is followed by inactivation (Fig. 3*e*). This finding supports a coherent picture of HERG as a member of the voltage-activated K^+ channel family, with an unusually slow activation and fast inactivation process. This may be different from the KAT1 K^+ channel (the other known inward rectifier with 6 transmembrane

segments), in which rectification is proposed to result from a voltage-activation mechanism with reversed polarity^{17–19}.

To explore further the relationship between HERG inactivation and *Shaker* C-type inactivation, we made mutations in the outer mouth of the HERG channel at position 631. This residue is homologous to position 449 in the outer mouth of *Shaker* channels, which has large effects on C-type inactivation^{16,20}. In HERG, mutating this position from serine to cysteine (S631C) produced a substantial increase in the rate of inactivation (Fig. 4*a*). Voltage-dependent activation and de-activation of the channel remained slow. By accident, our mutagenesis procedure also produced a double mutant (G628C:S631C) with very interesting properties: it lacked the normal HERG inactivation (Fig. 4*a*), and it lacked inward rectification (Fig. 4*b, c*). The instantaneous *I*-*V* reversed at a more positive voltage than the wild type, indicating a reduced selectivity for K^+ as is seen for mutations at the homologous position in *Shaker*²¹.

The single mutant G628C did not express, and we do not know the mechanism by which the two mutations in the double mutant interact to produce a functional channel that lacks inactivation. Nevertheless, the dramatic effects of pore mutations on the HERG inactivation process are consistent with the idea that it may be mechanistically similar to C-type inactivation of *Shaker* channels. The HERG inactivation process seems to be much faster and more voltage-dependent than *Shaker* C-type inactivation, but the rate and voltage dependence of *Shaker* inactivation can vary with mutation and ionic conditions^{20,22,23}. Thus it seems possible that the main differences between the two processes are quantitative and not qualitative.

The mechanism of rectification has important implications for the pharmacology of HERG and its *in vivo* equivalent (I_{Kr}), as blockers acting at the inner and outer mouths of the channel will interact differently with the inactivation process. Also, the kinetics of HERG inactivation indicate that these channels may be important in the normal physiological suppression of arrhythmias, specifically in suppressing the generation of premature afterbeats. The conditions under which we produced large outward (inhibitory) currents through these channels, prompt re-activation (Fig. 1*c*), are like those that occur during generation of a premature beat. Such a role for HERG in suppressing extra beats might help to explain the increased incidence of cardiac sudden death in patients that lack HERG currents, either because they carry a genetic defect (familial long QT syndrome) or because they are being treated with class III antiarrhythmics that block HERG channels. □

Received 24 November 1995; accepted 4 January 1996.

- Curran, M. E. et al. *Cell* **80**, 795–803 (1995).
- Warmke, J. W. & Ganetzky, B. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3438–3442 (1994).
- Sanguinetti, M. C., Jiang, C., Curran, M. E. & Keating, M. T. *Cell* **81**, 299–307 (1995).
- Trudeau, M. C., Warmke, J. W., Ganetzky, B. & Robertson, G. A. *Science* **269**, 92–95 (1995).
- Sanguinetti, M. C. & Jurkiewicz, N. K. *J. gen. Physiol.* **96**, 195–215 (1990).
- Shibasaki, T. *J. Physiol., Lond.* **387**, 227–250 (1987).
- Vandenberg, C. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2560–2564 (1987).
- Lu, Z. & MacKinnon, R. *Nature* **371**, 243–246 (1994).
- Nichols, C. G., Ho, K. & Hebert, S. J. *Physiol., Lond.* **476**, 399–409 (1994).
- Stanfield, P. R. et al. *J. Physiol., Lond.* **475**, 1–7 (1994).
- Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Science* **250**, 533–538 (1990).
- Zagotta, W. N., Hoshi, T. & Aldrich, R. W. *Science* **250**, 568–571 (1990).
- Demo, S. D. & Yellen, G. *Neuron* **7**, 743–753 (1991).
- Hoshi, T., Zagotta, W. N. & Aldrich, R. W. *Neuron* **7**, 547–556 (1991).
- Choi, K. L., Aldrich, R. W. & Yellen, G. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5092–5095 (1991).
- Yellen, G., Sodickson, D., Chen, T.-Y. & Jurman, M. E. *Biophys. J.* **66**, 1068–1075 (1994).
- Anderson, J. A., Huprikar, S. S., Kochian, L. V., Lucas, W. J. & Gaber, R. F. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3736–3740 (1992).
- Cao, Y., Crawford, N. M. & Schroeder, J. I. *J. biol. Chem.* **270**, 17697–17701 (1995).
- Hoshi, T. *J. gen. Physiol.* **105**, 309–328 (1995).
- López-Barneo, J., Hoshi, T., Heinemann, S. H. & Aldrich, R. W. *Recept. Chan.* **1**, 61–71 (1993).
- Heginbotham, L., Lu, Z., Abramson, T. & MacKinnon, R. *Biophys. J.* **66**, 1061–1067 (1994).
- Baukrowitz, T. & Yellen, G. *Neuron* **15**, 951–960 (1995).
- Baukrowitz, T. & Yellen, G. *Science* **271**, 653–656 (1996).
- Jurman, M. E., Boland, L. M., Liu, Y. & Yellen, G. *BioTechniques* **17**, 876–881 (1994).

ACKNOWLEDGEMENTS. We thank M. Jurman for assistance in preparing the mutants and the transfected cells, and G. Robertson for the cDNA clone for HERG. This work was partially supported by the National Institute of Neurological Diseases and Stroke (to G.Y.), by a Gottlieb Daimler-Karl Benz stipend (T.B.), and by an NIH training grant (P.L.S.).

Differential effects of the Rac GTPase on Purkinje cell axons and dendritic trunks and spines

Liqun Luo, Takao K. Hensch*, Larry Ackerman, Sandra Barbel, Lily Y. Jan & Yuh Nung Jan

Howard Hughes Medical Institute, Departments of Physiology, and Biochemistry and Biophysics, * Neuroscience Graduate Program, University of California at San Francisco, San Francisco, California 94143, USA

NEURONS contain distinct compartments including dendrites, dendritic spines, axons and synaptic terminals¹. The molecular mechanisms that generate and distinguish these compartments, although largely unknown, may involve the small GTPases Rac and Cdc42 (ref. 2), which appear to regulate actin polymerization³. Having shown that perturbations of Rac1 activity block the growth of axons but not dendrites of *Drosophila* neurons², we investigated whether this also applies to mammals by examining transgenic mice expressing constitutively active human Rac1 in Purkinje cells. We found that these mice were ataxic and had a reduction of Purkinje-cell axon terminals in the deep cerebellar nuclei, whereas the dendritic trees grew to normal height and branched extensively. Unexpectedly, the dendritic spines of Purkinje cells in developing and mature cerebella were much reduced in size but increased in number. These 'mini' spines often form supernumerary synapses. These differential effects of perturbing Rac1 activity indicate that there may be distinct

mechanisms for the elaboration of axons, dendrites and dendritic spines.

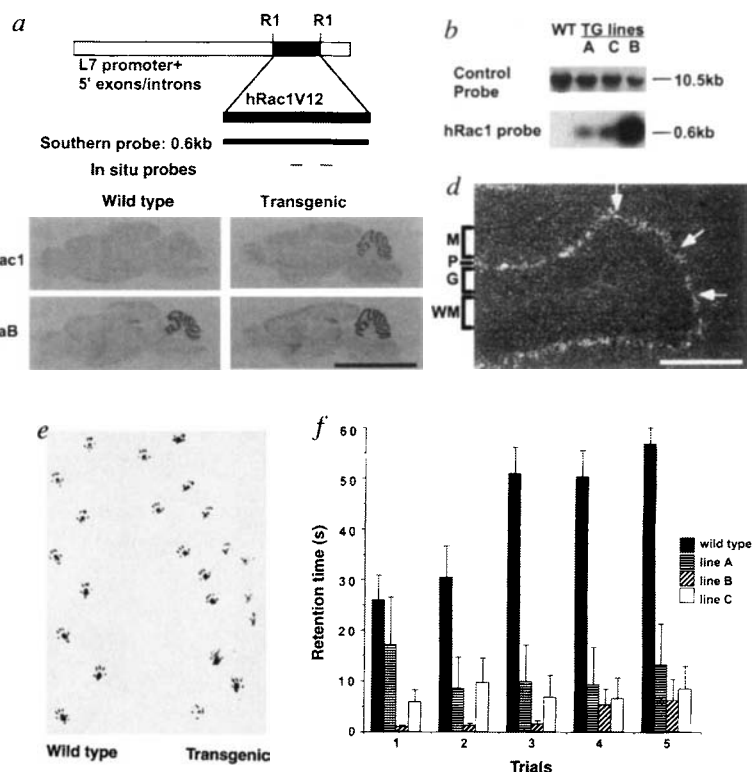
We chose to examine the role of Rac in the morphogenesis of the mouse cerebellar Purkinje cell because of its clear axon versus dendrite polarity¹, its well characterized anatomy and synaptic connections⁴, and its non-vital yet important function in motor coordination⁵. We used the Purkinje cell-specific L7 promoter^{6,7} to express human Rac1 with a point mutation V12 (hRac1V12) that renders it constitutively active^{2,3} (Fig. 1a). We selected the V12 mutation because in *Drosophila* it disrupts axon outgrowth in a manner qualitatively similar to that of a dominant negative mutation but yields a stronger phenotype². Three independent transgenic lines were generated (Fig. 1b). The transgene expression in the brain is limited to cerebellar Purkinje cells (Fig. 1c, d). All three lines showed similar behavioural and anatomical defects as described below, indicating that these defects resulted from hRac1V12 expression in Purkinje cells.

Transgenic mice were ataxic from the time of weaning. They trembled at the onset of motion, did not walk in straight lines (Fig. 1e), and performed poorly on the rotarod assay⁸ (Fig. 1f). These defects are indicative of cerebellar dysfunction⁹ and were not obviously age-dependent. None of the transgenic mice died before 15 months of age.

The position and density of Purkinje cells in transgenic mice appeared normal as revealed by Nissl staining (not shown) or antibody staining for the Purkinje cell markers calbindin⁹ (Fig. 2a–d) and the inositol trisphosphate (InsP₃) receptor¹⁰ (not shown). The dendritic trees retained a large degree of complexity (Fig. 2d). However, immunocytochemical staining for calbindin, a cytoplasmic protein, the InsP₃ receptor, an endoplasmic reticulum protein, and a synaptic vesicle protein synaptophysin¹¹ (Fig. 2e and

FIG. 1 Generation and behavioural characterization of transgenic mice expressing constitutively active human Rac1 in Purkinje cells. **a**, The transgene construct, with the probes used for Southern and *in situ* hybridization indicated. R1, *Eco*RI. **b**, Three transgenic (TG) lines A, B, and C were verified on a genomic Southern blot by the presence of a 0.6-kb *Eco*RI fragment which was absent in wild-type (WT) littermates. Copy numbers were estimated to be 2, 50 and 3 for A, B and C, respectively, after comparison with a control probe for a single-copy gene using PhosphorImager quantitation of the Southern blot. No strict correlation of the copy number with severity of phenotype was observed. **c**, Adjacent sagittal sections of a transgenic mouse or a WT littermate were treated with transgene-specific probes complementary to the hRac1 sequence (nucleotides (nt) 304–348 and 415–459)²¹ or a probe for a Purkinje-cell marker probe calbindin (CaB) (nt 175–219)²², showing cerebellum-specific expression of hRac1 in transgenic mice. **d**, High-magnification micrograph reveals the hRac1 hybridization signals in the Purkinje cell layer. M, P, G, WM indicate molecular, Purkinje, granular layers and white matter, respectively. The Purkinje cell layer is also indicated by arrows. **e**, Hind footprint assay. Instead of walking along a straight line as do WT littermates, transgenic mice wobbled, with a wider base and shorter steps, from side to side on a 8" × 14" longitudinally folded sheet of paper. **f**, Compared with WT, transgenic mice from all three lines performed poorly on the rotarod assay⁸, which measures the duration over which mice remain standing on a 6-cm diameter rod rotating at 4 r.p.m. Five consecutive trials were given for a maximum of 60 s. The numbers of animals tested were 14, 8, 16 and 14 for WT, TG lines A, B and C, respectively. Data shown are from progeny of a single F₁ transgenic mouse for each line. Error bars indicate s.e.m. Scale bar: 5.6 mm in c, 300 μm in d.

METHODS. A glycine-12 to valine-12 mutation was introduced by oligo-directed mutagenesis of hRac1, which encodes a protein with amino-acid sequence identical to that of mouse Rac1 (ref. 23). The hRac1V12 open reading frame was subcloned into the *Bam*HI site of the L7 promoter vector⁷ and used for transgenic mice production²⁴. Hybrids of C57BL/6J and DBA2



(B6D2) were used as hosts for transgenes. Mice used in this study were progeny of the founders or F₁ TG crossed to B6D2 or C57BL/6J for heterozygotes, or each other for homozygotes (used only in Fig. 4h). *In situ* hybridization was done on 5-μm cryostat sections²⁵. X-ray films (c) or emulsions (d) were developed 9 or 40 days later.

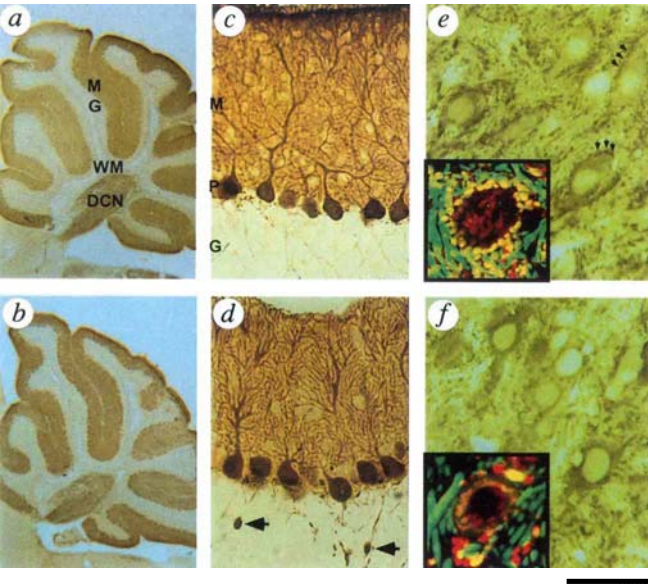
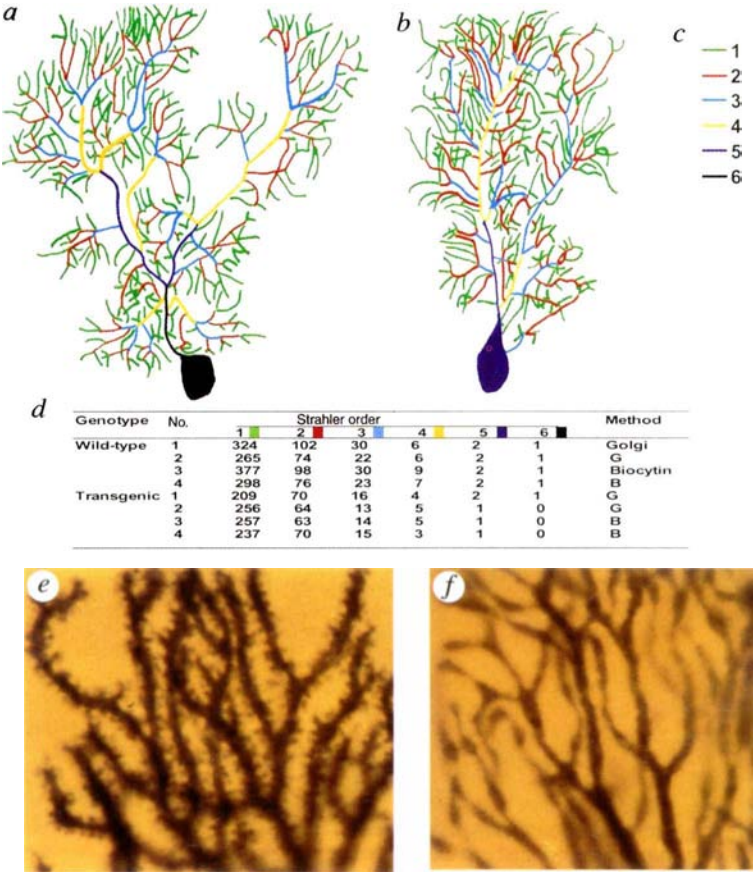


FIG. 2 Cerebellar histology. *a–d*, Anti-calbindin staining²⁶ of sagittal sections reveals similar organization of the cerebellar cortex (abbreviations as for Fig. 1*d*), the deep cerebellar nuclei (DCN) (*a*, WT; *b*, TG), and similar Purkinje cell density, dendritic trees and proximal axons in WT (*c*) and TG (*d*). Arrows in *d* indicate abnormal torpedo-like axon swellings^{12,27} that are frequently found in transgenic mice. Sometimes there were no detectable axons distal to the torpedoes. *e, f*, Sagittal sections of medial DCN stained with calbindin. Whereas calbindin-immunoreactive axons (parallel or perpendicular to the section plane) are still present in transgenic mice (*f*), there is a drastic reduction of Purkinje-cell axon terminals around the cell bodies of the DCN neurons compared with wild type (arrows in *e*). Insets of *e* and *f* show an optical section of representative medial DCN neurons double-labelled for calbindin (green) and synaptophysin (red). In wild-type, the cell body is surrounded by double-labelled (yellow) Purkinje cell axon terminals (*e*). A previous study²⁸ estimated that 70–80% of the synapses onto the DCN cell bodies were from Purkinje cells. In transgenic mice, DCN neurons are surrounded by fewer structures double-labelled with synaptophysin and calbindin, although synaptophysin within DCN neurons (presumably made by themselves) remains detectable (*f*). There is some variation for the extent of Purkinje-cell terminal loss across different regions of the DCN; the medial part was affected most severely (data not shown). This is consistent with the notion that the timing of transgene expression relative to axonogenesis is important, because L7 expression starts in the vermal Purkinje cells⁶, which project to medial DCN. Scale bar in *a, b*, 1 mm; in *c, d*, 66 μ m; in *e, f*, 42 μ m; and in insets of *e* and *f*, 25 μ m.

METHODS. Antibody dilutions were: mouse monoclonal anti-calbindin (Sigma), 1:10,000; rabbit anti-synaptophysin¹¹, 1:250; rabbit anti-InsP₃ receptor¹⁰, 1:250; DTAF- or Cy3-conjugated secondary (Jackson), 1:200 or 1:1,000, respectively. Images were collected on a Nikon Optiphot or a BioRad MRC 600 confocal microscope (double labelling).

FIG. 3 Dendritic tree morphology. *Camera lucida* drawings of individual Purkinje cells (*a*, WT; *b*, TG) after filling intracellularly with dye reveal the complexity of their dendritic trees. The colour bars (in *c*) represent the Strahler order for quantitation of the complexity¹⁵. Terminal branches are defined as 1. When two 'n' order branches meet, an 'n + 1' order is created, whereas an 'n + 1' order meeting with an 'n' remains 'n + 1'. The dendritic tree in transgenic mice appears narrower than it actually is because it lies in a slanted plane relative to the plane of section. *d*, Quantitation of the dendritic tree complexity by the Strahler method¹⁵. B, biocytin injection method; G, Golgi method. The dendritic tree complexity is slightly but significantly reduced in transgenic mice ($P = 0.03$ for the terminal branches, one-tailed Students *t*-test). High-magnification micrographs of portions of biocytin-filled Purkinje cell dendritic trees from a wild-type (*e*) and a transgenic mouse (*f*) reveal the apparent loss of dendritic spines in transgenic mice at this resolution. Scale bar: 38 μ m for *a* and *b*; 8.6 μ m for *e* and *f*.

METHODS. For dye-filling of individual Purkinje cells, 500- μ m sagittal vermal slices were prepared and maintained as described²⁹. Whole-cell patch clamp recordings were obtained from Purkinje cells with a patch electrode (3–8 M Ω) in the whole-cell current clamp mode (Axoclamp-2B). The pipette solution contained (in mM): 122.5 potassium gluconate, 17.5 KCl, 10 HEPES buffer, 0.2 EGTA, 8 NaCl, 2.0 Mg-ATP, 0.3 Na₃-GTP, and 0.15% Biocytin (pH 7.2, 290–300 mOsm). Both wild-type and transgenic Purkinje cells displayed similarly healthy resting potential (lower than –55mV) and input resistance (> 100 M Ω). After dialysing the cell contents for a minimum of 45 min, slices were fixed, resectioned at 50 μ m, and processed according to standard avidin–biotin complex techniques (Vector), with the biocytin label being visualized using the nickel-intensified diaminobenzidine reaction. Golgi–Colonnier staining was done as described³⁰; after the last AgNO₃ reaction, tissues were incubated in 70% ethanol for 1 h and sectioned further with a vibratome at 100 μ m. Individual Purkinje cells were traced on a Nikon microscope with a 40x oil lens.



f, insets) revealed a drastic reduction of Purkinje-cell axon terminals in the deep cerebellar nuclei (DCN) of transgenic mice (Fig. 2*f*) as compared with wild type (Fig. 2*e*, arrows). A similar reduction of Purkinje cell terminals was found in young adult (1 month) and in older mice (6 months). We also observed unusual swelling of Purkinje axons in the granular layer (Fig. 2*d*), reminiscent of 'torpedoes' found in 'hyperspiny Purkinje cell' (hpc) mutant mice¹² or in mice subjected to Purkinje cell axotomy¹³.

In *Drosophila* neurons, the timing of Drac1V12 expression relative to that of axonogenesis determines the phenotype: early expression blocks axon initiation, whereas later expression arrests axon elongation². The L7 promoter becomes active on embryonic day 17 for a small subset of vermal Purkinje cells and around

postnatal days 2–7 for most Purkinje cells⁶, whereas the onset of Purkinje cell axonogenesis takes place earlier (before birth)¹⁴. This relatively late onset of hRac1V12 expression could account for the presence of residual axons and the loss of Purkinje cell axon terminals.

In contrast to axonogenesis, growth of the Purkinje cell dendritic tree occurs mostly during the second postnatal week^{14,15}, after the onset of L7 expression. However, hRac1V12 expression did not appear to perturb the vertical growth of the Purkinje cell dendritic trees, as revealed by the extent of calbindin staining in the molecular layer (Fig. 2*a–d*). We further studied the morphology of individual Purkinje cells filled with biocytin or stained by the Golgi method (Fig. 3). Purkinje cell dendritic trees in transgenic mice retained a normal planar organization¹. The branching complexity, as quantitated by the Strahler method, was only slightly reduced compared to wild type (Fig. 3*d*). However, the dendritic spines, a characteristic feature of Purkinje cell dendritic trees (Fig. 3*e*), were strikingly different in transgenic mice. Spines were hardly visible under the light microscope in dye-filled (Fig. 3*f*) or Golgi-stained (not shown) Purkinje cells.

The fine structure of dendritic spines was further examined in Golgi-impregnated Purkinje cells with a 300 kV electron microscope. Compared to the dendritic spines in wild-type animals, which were well spaced and composed of thin necks and large heads (Fig. 4*a*), the dendritic spines in transgenic mice appeared much smaller and more numerous (Fig. 4*b*). This abnormality was already evident during postnatal day 9 to 15, when spines in wild-type animals are still in the process of acquiring adult number and morphology (Fig. 4*c, d*). Thin-section electron microscopy confirmed that transgenic mice contained abnormally small Purkinje cell dendritic spines⁴ (asterisks in Fig. 4*e* and *f*) bearing synapses with parallel fibres and climbing fibres. The average cross-sectional area of spines with synapses was 0.17–0.21 μm^2 in wild-type animals (Fig. 4*g*). This was reduced to 0.075–0.09 μm^2 in animals heterozygous for the transgene and even smaller in homozygous animals (0.051–0.053 μm^2). Interestingly, whereas spines with double synapses were rare in normal animals (Fig. 4*h*), they were found frequently in every transgenic animal (Fig. 4*f*, double asterisks; Fig. 4*h*). Occasionally, we encountered

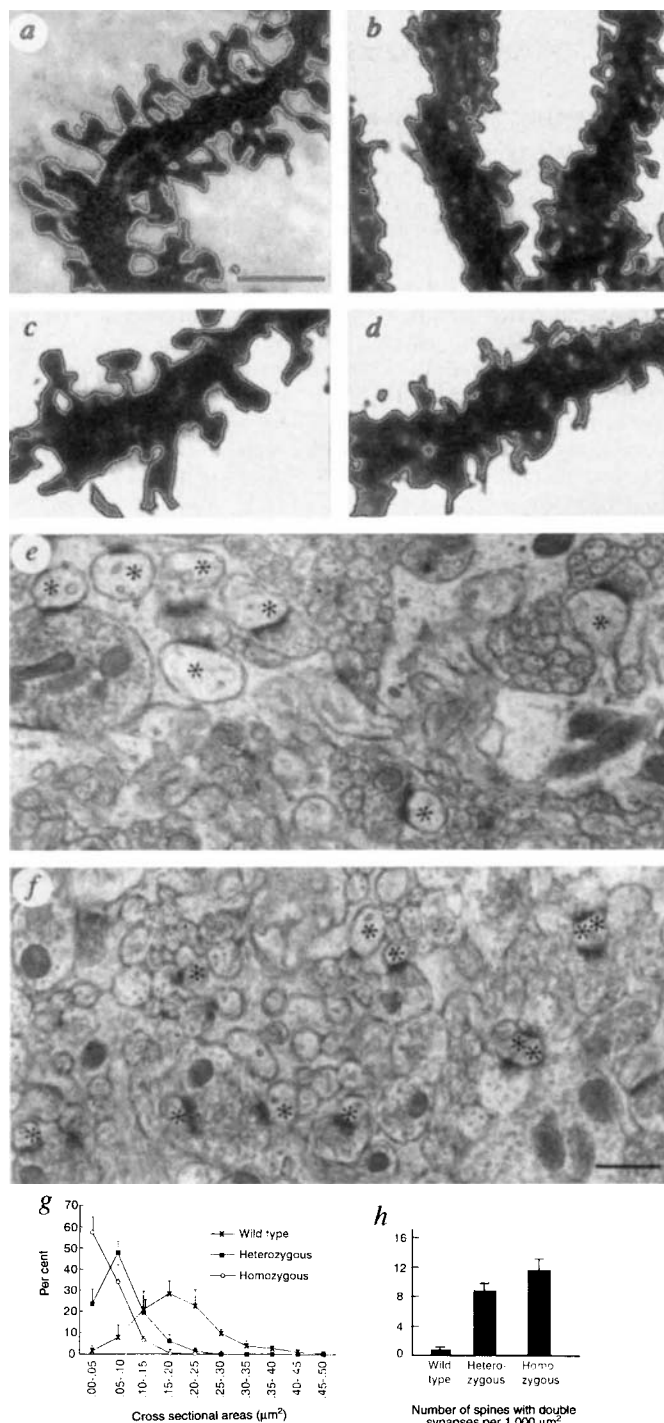


FIG. 4 Ultrastructural analyses. Representative 300 kV electron micrographs of Golgi-impregnated Purkinje cell dendrites (*a*, WT; *b*, TG at 2 months) reveal that in transgenic mice spines are smaller and more numerous. These phenotypes were evident at P12 (*c*, WT; *d*, TG), when wild-type animals are in the process of generating new spines and spine heads. Representative thin-section electron micrographs (*e*, WT; *f*, TG at 6 months) reveal a reduction of cross-sectional areas of transgenic Purkinje cell dendritic spines (asterisk). Also abnormal is the frequent occurrence of spines that synapse with two presynaptic terminals (double asterisks). These parameters are quantified in a histogram of the distribution of cross-sectional areas of spines with synapses in *g*, and of the number of double synapses per unit area in *h*. For both *g* and *h*, thin-section electron micrographs covering 700–1,000 μm^2 neuropil regions from each mouse were used for quantitation. The number of animals examined was 4, 4 and 2 for wild-type, heterozygous and homozygous, respectively, with ages between 1.5 to 8 months. Error bars indicate standard deviations. Scale: *a–d*, 2 μm ; *e, f*, 0.5 μm .

METHODS. For 300 kV electron microscopy, 2- μm sections of Golgi-impregnated Purkinje cells were imaged using a Philips 430 electron microscope with a CCD system. For thin-section electron microscopy, mice were perfused with 4% formaldehyde (EM grade), 2% glutaraldehyde in 0.1 M phosphate buffer (pH 7.2), and processed for EM according to standard methods⁴. 100-nm-thin sections of over 700 μm^2 random and non-overlapping neuropil regions of the molecular layer from vermal mid-folium V were photographed with a Philips 400 electron microscope. Spines with synapses were circled on a transparency, scanned into a computer, and the enclosed areas calculated by a program AREASTAT (written by E. Frise of UCSF).

two juxtaposed presynaptic terminals facing the same side of a spine, suggesting that at least some of the double synapses were contributed by two distinct presynaptic terminals, in contrast to the one synapse per spine observed in wild-type Purkinje cells¹⁶.

The evidence presented here indicates that Rac affects axons and presynaptic terminals but does not affect dendritic trunk in mouse Purkinje cells, analogous to the effects of Drac1 on axons but not dendrites of *Drosophila* neurons. Surprisingly, dendritic spine morphogenesis is severely altered. This phenotype is found in developing as well as mature cerebella, and is unlikely to be a secondary consequence of Purkinje axon damage^{12,13}, suggesting that the hRac1V12 transgene directly affects the formation of dendritic spines during development. This does not exclude the possibility that Rac also affects the maintenance of spine morphology in adult. The Rac/Rho/Cdc42 subfamily of GTPases have two distinct functions: to regulate the actin cytoskeleton^{3,17} and to transduce signals from the membrane to the nucleus¹⁷. It is possible that expression of activated Rac affects the signalling events necessary for Purkinje cell differentiation, thus preventing the elaboration of spines and presynaptic terminals. However, the phenotypes in Rac transgenic mice do not correspond to any normal developmental stages of Purkinje cells and therefore do not represent a simple developmental arrest. Moreover, similar expression of Drac1V12 in *Drosophila* neurons blocked axon initiation and elongation without changing neuronal fate². Thus, an explanation for the Rac phenotype in both mouse and fly could be a specific involvement of Rac in the formation of particular neuronal processes. The differential effects of Rac may therefore reveal mechanistic differences underlying the elaboration of axons, dendritic trunks and dendritic spines. It is important to note, however, that studies of loss-of-function mutations are required to verify whether Rac1 is normally involved in spine formation.

Since their discovery¹, dendritic spines have been proposed to be important in functions such as increasing postsynaptic surface area or creating microdomains for synaptic integration and memory storage¹⁸. How dendritic spines form is not known, although in Purkinje cells it is likely to involve intrinsic mechanisms¹⁹ as well as inductive signals²⁰. Our experiments provide a potential molecular handle to further examine the intrinsic mechanisms of spine formation and perhaps their regulation by extrinsic factors. The possibility of altering spine size and number in transgenic mice also presents an opportunity to test the physiological functions of dendritic spines. □

Received 20 September 1995; accepted 15 January 1996.

1. Cajal, S. R. *Histology of the Nervous System of Man and Vertebrates* (Oxford Univ. Press, 1995).
2. Luo, L., Liao, Y. J., Jan, L. Y. & Jan, Y. N. *Genes Dev.* **8**, 1787–1802 (1994).
3. Hall, A. A. *Rev. Cell Biol.* **10**, 31–54 (1994).
4. Palay, S. L. & Chan-Palay, V. *Cerebellar Cortex: Cytology and Organization* (Springer, New York, 1974).
5. Ito, M. *The Cerebellum and Neural Control* (Raven, New York, 1984).
6. Oberdick, J. et al. *Neuron* **10**, 1007–1018 (1993).
7. Smeyne, R. J. et al. *Molec. cell. Neurosci.* **6**, 230–251 (1995).
8. Dunham, N. W. & Miya, T. S. *J. Am. Pharmac. Assoc.* **46**, 208–209 (1957).
9. Jande, S. S., Maler, L. & Lawson, D. E. M. *Nature* **294**, 765–767 (1981).
10. Mignery, G. A., Südhof, T. C., Takei, K. & De Camilli, P. *Nature* **342**, 192–195 (1989).
11. Jahn, R., Schiebler, R., Quimet, C. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4137–4141 (1985).
12. Sotelo, C. *J. Neurocytol.* **19**, 737–755 (1990).
13. Dussart, I. & Sotelo, C. *J. comp. Neurol.* **347**, 211–232 (1994).
14. Altman, J. *J. comp. Neurol.* **145**, 399–464 (1972).
15. Berry, M. & Bradley, P. *Brain Res.* **112**, 1–35 (1976).
16. Harris, K. M. & Stevens, J. K. *J. Neurosci.* **8**, 4455–4469 (1988).
17. Vojtek, A. B. & Cooper, J. A. *Cell* **82**, 527–529 (1995).
18. Harris, K. W. & Kater, S. B. *A. Rev. Neurosci.* **17**, 341–371 (1994).
19. Rakic, P. & Sidman, R. L. *J. comp. Neurol.* **152**, 133–162 (1973).
20. Baptista, C. A., Hatten, M. E., Blazeski, R. & Mason, C. A. *Neuron* **12**, 243–260 (1994).
21. Didsbury, J., Weber, R. F., Bokoch, G. M., Evans, T. & Snyderman, R. *J. biol. Chem.* **264**, 16378–16382 (1989).
22. Wood, T. L. et al. *DNA* **7**, 585–593 (1988).
23. Moll, J., Sansig, G., Fattori, E. & van der Putten, H. *Oncogene* **6**, 863–866 (1991).
24. Hogan, B., Constantini, F. & Lacy, E. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory, Cold Spring Harbor, 1986).
25. Widsen, W. & Morris, B. J. in *In Situ Hybridization Protocols for the Brain* (eds Widsen, W. & Morris, B. J.) 9–34 (Academic, London, 1994).
26. Jones, K. R., Farinas, I., Backus, C. & Reichardt, L. F. *Cell* **76**, 989–999 (1994).
27. Guenet, J.-L., Sotelo, C. & Mariani, J. *J. Hered.* **74**, 105–108 (1983).

28. Roffler-Tarlov, S., Beart, P. M., O'Gorman, S. & Sidman, R. L. *Brain Res.* **168**, 75–95 (1979).
29. Ajima, A., Hensch, T., Kado, R. T. & Ito, M. *Neurosci. Res.* **12**, 281–286 (1991).
30. Fairen, A. & Smith-Fernandez, A. *Microsc. Res. Tech.* **23**, 289–305 (1992).

ACKNOWLEDGEMENTS. We thank J. Oberdick for the L7 promoter; R. Qiu and M. Symons for hRac1 cDNA; P. De Camilli and R. Jahn for antibodies; M. Stryker for the use of the brain-slice physiology setup (supported by an HFSP Grant); D. Agard and J. Sedat for the use of electron microscopes; M.-M. Jiang, D. Merlis and Y. Ruan for technical assistance; E. Frise for a computer program to quantify spine areas and help with artwork; and I. Farinas, K. Harris, I. Herskowitz, J. Knoblich, X. C. Liao, D. Stretavan, B. Zhao and W. Zhong for discussion and for comments on the manuscript. L.L. was a postdoctoral fellow of the Jane Coffin Childs Fund for Biomedical Research. T.K.H. is a Howard Hughes predoctoral fellow in the UCSF Neuroscience Graduate Program. L.Y.J. and Y.N.J. are investigators of the Howard Hughes Medical Institute. This work was supported by a NIMH Silvio Conte Center for Neuroscience Research Grant at UCSF.

Insulin action impaired by deficiency of the G-protein subunit $G_{i\alpha 2}$

Christopher M. Moxham & Craig C. Malbon

Department of Molecular Pharmacology, Diabetes and Metabolic Diseases Research Program, University Medical Center, SUNY/Stony Brook, Stony Brook, New York 11794-8651, USA

INTEGRATION of information between tyrosine kinase¹ and G-protein-mediated pathways² is necessary, but remains poorly understood. Here we use cells from transgenic mice harbouring inducible expression of RNA antisense to the gene encoding $G_{i\alpha 2}$ (refs 3, 4) to show that $G_{i\alpha 2}$ is critical for insulin action. $G_{i\alpha 2}$ deficiency in adipose tissue and liver produces hyperinsulinaemia, impaired glucose tolerance and resistance to insulin *in vivo*. Insulin resistance affects glucose-transporter activity and recruitment, counterregulation of lipolysis, and activation of glycogen synthase, all of which are cardinal responses to insulin⁵. $G_{i\alpha 2}$ deficiency increases protein-tyrosine phosphatase activity and attenuates insulin-stimulated tyrosine phosphorylation of IRS (insulin-receptor substrate 1) *in vivo*. $G_{i\alpha 2}$ deficiency creates a model for the insulin resistance characteristic of non-insulin-dependent diabetes mellitus (NIDDM)⁶, implicating $G_{i\alpha 2}$ as a positive regulator of insulin action.

The phenotype of BDF1 mice harbouring the pPCK-ASG_{iα2} transgene and $G_{i\alpha 2}$ -deficient tissues (>90% suppression) is runt^{3,4}, which prompted us to examine the metabolic consequences of the $G_{i\alpha 2}$ deficiency. Fasting serum-insulin levels are elevated in transgenic mice compared with controls (0.46 ± 0.03 and 0.28 ± 0.01 ng ml⁻¹, $n = 6$, respectively). Fasting blood-glucose levels, in contrast, are equivalent in control and transgenic mice (55 ± 5 and 58 ± 4 mg dl⁻¹, $n = 6$, respectively). After receiving a single bolus of glucose, the glucose tolerance of the transgenic mice was impaired (Fig. 1a). After loading, serum glucose levels decline to <80 mg dl⁻¹ within 30 min for control littermates, requiring 60–120 min in transgenic mice. The results of an insulin-tolerance test buttress those of the glucose-tolerance test. Insulin induces a rapid and sustained hypoglycaemia in fasted, control mice, but not in pPCK-ASG_{iα2} transgenic mice (Fig. 1b). The runt phenotype^{3,4}, fasting hyperinsulinaemia, and impaired glucose and insulin tolerance of the pPCK-ASG_{iα2} transgenic mice all resemble the phenotype of insulin resistance characteristic of NIDDM⁶.

Insulin counterregulates the effects of catecholamines on lipolysis^{5,6}. At a maximal concentration of the β-adrenergic agonist isoprenaline, insulin inhibits lipolysis in adipocytes from control mice (Fig. 2a) but this counterregulatory effect of insulin is impaired in adipocytes of pPCK-ASG_{iα2} mice. At 0.1 nM insulin, adipocytes from transgenic mice were not counterregulated, whereas the response was maximal for those from control littermates. To what extent, if any, does the increased cyclic AMP in $G_{i\alpha 2}$ -deficient adipocytes (Fig. 2b) contribute to insulin resistance?

We measured lipolysis in adipocytes from BDF1 control mice stimulated with a concentration of isoprenaline that mimics that of intracellular cAMP in $G_{i\alpha 2}$ -deficient animals. Even at equivalent intracellular cAMP concentrations, adipocytes from control mice show markedly greater inhibition of lipolysis in response to insulin than do $G_{i\alpha 2}$ -deficient adipocytes (Fig. 2c). Differences in intracellular cAMP accompanying $G_{i\alpha 2}$ deficiency cannot account for the profound insulin resistance.

Insulin promotes glycogen deposition in liver and skeletal muscle by converting glycogen synthase from a glucose-6-phosphate (G6P)-dependent (D-form) to an independent enzyme (I-form) which is more active at physiological concentrations of G6P⁷. In control mice, insulin stimulates a rapid increase in the amount of I-form in liver (Fig. 2d) and skeletal muscle (Fig. 2e), reaching a peak at 3 min in liver and 2 min in skeletal muscle. For transgenic mice, insulin failed to stimulate glycogen synthase activity in both liver (Fig. 2d) and skeletal muscle (Fig. 2e). Total synthase activity of liver and skeletal muscle was equivalent in control and transgenic mice.

Insulin promotes glucose uptake by adipocytes⁸ and 3-O-methyl-glucose uptake in adipocytes of control mice ~30-fold (Fig. 3a). $G_{i\alpha 2}$ -deficient adipocytes, in contrast, are insulin-resistant. We found that basal glucose transport rates were equivalent in control and transgenic mouse adipocytes (0.12 ± 0.03 and 0.08 ± 0.02 fmol per cell per min, $n = 6$, respectively). Any contribution of raised intracellular cAMP to insulin resistance and hence glucose transport by $G_{i\alpha 2}$ -deficient adipocytes was established by stimulating adipocytes from control animals with sufficient β -adrenergic agonist to 'clamp' their cAMP at levels of unstimulated, $G_{i\alpha 2}$ -deficient adipocytes. At equivalent levels of intracellular cAMP (Fig. 2b), adipocytes isolated from control mice still show a robust (18-fold) increase in glucose transport activity.

We tested insulin's ability *in vivo* to promote translocation of the glucose transporter GLUT4 to the plasma membrane of adipocytes⁹. Insulin administration *in vivo* stimulates a time-dependent translocation of GLUT4 from the low-density microsomal (L) to the plasma membrane (P) fraction of adipose tissue

of BDF1 control mice within 6 min, reaching a maximum by 12 min (Fig. 3b). For pPCK-ASG $_{i\alpha 2}$ transgenic mice, however, insulin administration failed to promote recruitment of GLUT4 to the plasma-membrane-enriched fraction at 6, 9 or 12 min. Probing blots of the same fractions with antibodies against the insulin receptor β -subunit showed that sample loading was equivalent (Fig. 3b).

The whole-body measurements of glucose and insulin tolerance, together with measurements of glucose transport activity in adipocytes, glycogen synthase activity in liver and skeletal muscle, and assay of GLUT4 recruitment provide compelling evidence for profound insulin resistance in $G_{i\alpha 2}$ -deficient target tissues.

The insulin resistance of $G_{i\alpha 2}$ -deficient tissues suggests a lesion(s) near the insulin receptor, common to several insulin-sensitive pathways⁵⁻⁹. Insulin administration *in vivo* yields a rapid increase in phosphotyrosine phosphorylation of IRS-1 (Fig. 4a), a proximal element for insulin signalling⁵. Insulin-stimulated tyrosine phosphorylation of IRS-1 is absent in $G_{i\alpha 2}$ -deficient adipose tissue (Fig. 4a), suggesting that either insulin-receptor kinase activity is reduced or tyrosine phosphatase activity is increased, or both. Tyrosine kinase activity of insulin receptor preparations, partially purified from liver of control and transgenic mice, was equivalent, both displaying ~3-fold increases in activity following insulin stimulation (not shown). Total cellular phosphotyrosine phosphatase (PTPase) activity in adipose, liver and skeletal muscle from the pPCK-ASG $_{i\alpha 2}$ transgenic mice was markedly increased compared to control littermates (Fig. 4b). PTPase activity in particulate (P) and cytosolic (C) fractions was increased in adipose, liver and skeletal muscle of transgenic mice. Insulin-resistant tissues of the transgenic mice show a 1–2-fold increase in PTP1B expression (Fig. 4c), with significant amounts of PTP1B present in the cytosolic fraction (Fig. 4c, d). PTP1D (also called PTP2C and *Syp*) is an SH2-domain-containing PTPase implicated in several, but not all, pathways of insulin action^{10,11}. Unlike PTP1B, expression of *Syp* was not altered significantly by $G_{i\alpha 2}$ deficiency (Fig. 4c). PTPase provides a critical link between $G_{i\alpha 2}$ deficiency and insulin resistance.

Our findings extend earlier reports of insulin resistance in

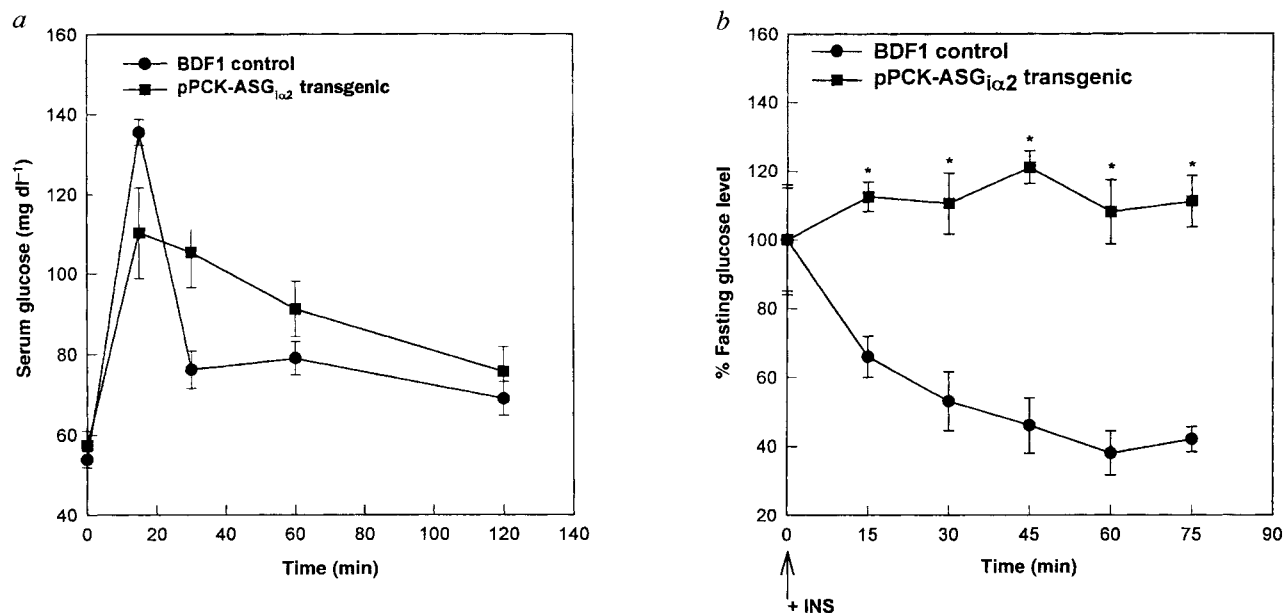


FIG. 1 pPCK-ASG $_{i\alpha 2}$ transgenic mice have increased insulin concentrations and impaired tolerance to glucose and insulin *in vivo*. Control and transgenic mice were made to fast overnight for 15 h. a, Amount of glucose in blood samples taken from mouse tails at the times indicated after glucose loading (2.5 mg per g body weight, injected i.p.). b, Glucose in mouse tail-blood samples at the indicated times after insulin injection (0.75 IU per kg body weight, injected i.p.). Results are the mean values $\pm$ s.e.m. from 6 mice for

each condition. Asterisks denote values statistically different from control ($P < 0.05$).

METHODS. Mice were maintained on a normal light/dark cycle, and their glucose metabolism and insulin levels analysed between 8:00 and 10:00 a.m. only. Glucose was determined using a One Touch II glucometer (Life Scan Technologies). All animals were handled in accordance with the guidelines established by the Institutional Animal Care and Use Committee at SUNY/Stony Brook.

FIG. 2 G_{i2} deficiency abolishes counterregulation of lipolysis and stimulation of glycogen synthase by insulin. *a*, Lipolysis (glycerol release) was measured in adipocytes isolated from BDF1 control (●) or pPCK- ASG_{i2} transgenic mice (■) after a 30-min challenge with insulin in the presence of a maximal concentration ($10 \mu M$) of the β -adrenergic agonist isoprenaline. *b*, The levels of cAMP in untreated, adipocytes of control mice (○) are normalized to those of untreated, G_{i2} -deficient adipocytes (■) after a 10-min challenge with 25 nM isoprenaline at $37^\circ C$ (●). *c*, Insulin markedly inhibits lipolysis in normal adipocytes (●) challenged with a submaximal dose of isoprenaline (25 nM), but not lipolysis in untreated, G_{i2} -deficient adipocytes (■). *d* and *e*, Insulin stimulates glycogen synthase activity in liver and skeletal muscle of control but not pPCK- ASG_{i2} transgenic mice.

METHODS. Adipocytes were acutely prepared from epididymal fat pads by digestion with collagenase (1 mg ml^{-1}) in Krebs-Ringer phosphate buffer supplemented with BSA (3% w/v) and adenosine deaminase (0.5 U ml^{-1})⁴. Determinations of cAMP accumulation⁴ or glycerol release²⁴ were performed as described. For determinations of glycogen synthase activity, control and transgenic mice (8–10 weeks) were anaesthetized with ketamine/xylazine solution ($120/80 \text{ mg per kg body weight}$, respectively, injected i.p.). Once deep anaesthesia was achieved, the abdominal cavity was opened and the viscera exposed. A bolus of porcine insulin (5 U) was injected into the inferior vena cava. At the indicated times after insulin injection, portions of the liver and hindlimb skeletal muscle were removed and measurements of glycogen synthase activity performed as described²⁵. Protein was determined as described²⁶. Results are mean values $\pm$ s.e.m. ($n = 6$). Asterisks denote values significantly different from control ($P < 0.05$). Basal levels of glycerol release were 5.32 ± 0.04 and $7.68 \pm 0.38 \text{ nmol per } 10^5 \text{ cells per } 30 \text{ min}$ for adipocytes from control and transgenic mice, respectively. Maximal glycerol release in response to $10 \mu M$ isoprenaline was 16.47 ± 0.62 and $17.13 \pm 0.9 \text{ nmol per } 10^5 \text{ cells per } 30 \text{ min}$ in adipocytes from control and transgenic mice, respectively.

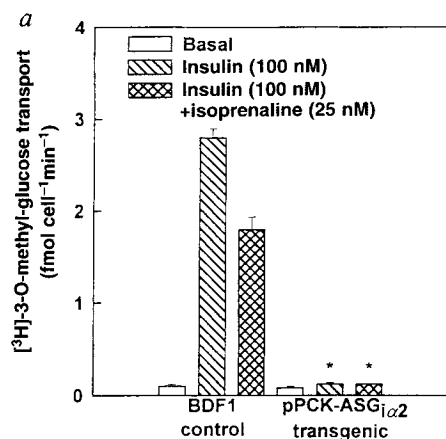
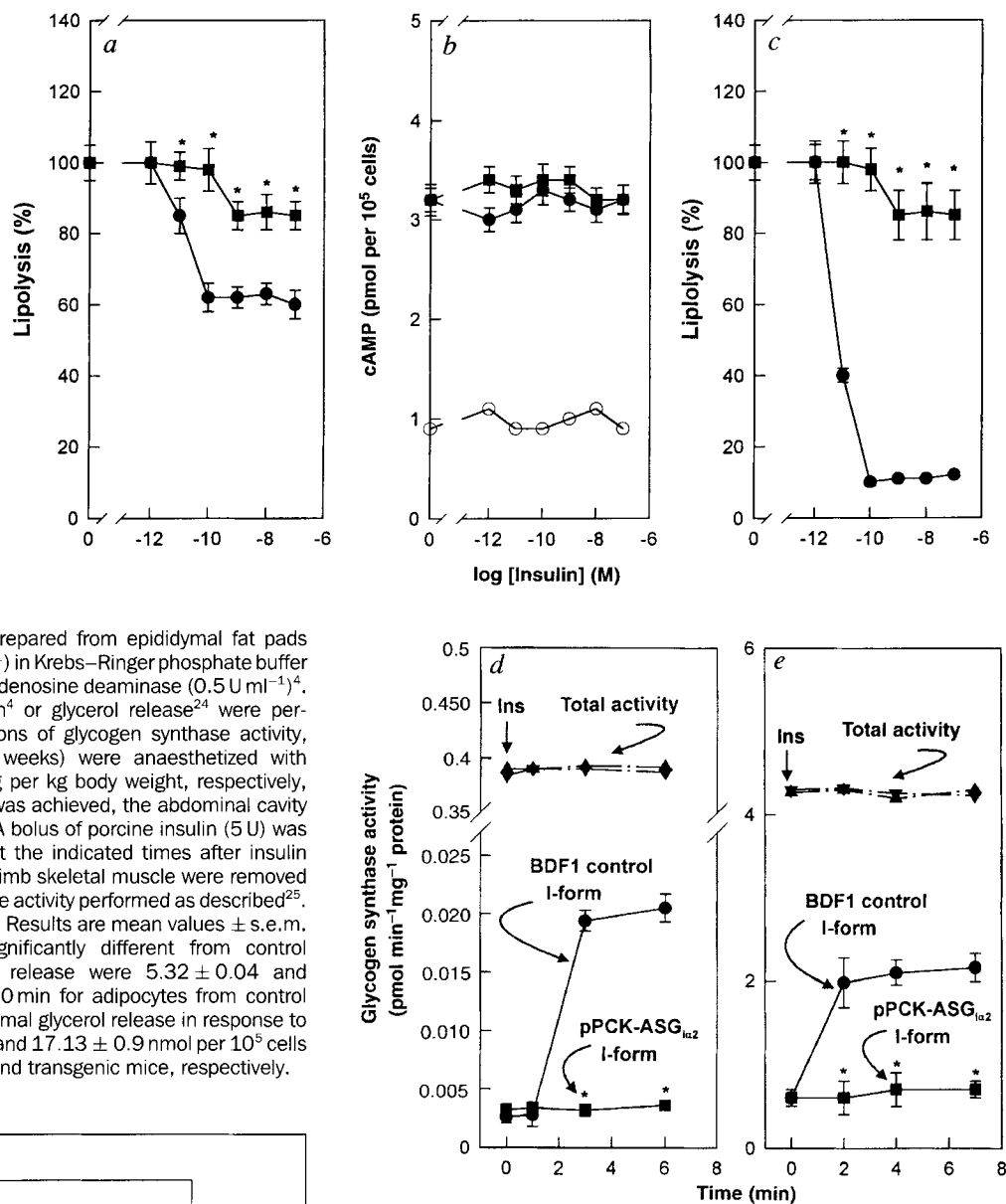


FIG. 3 G_{i2} deficiency abolishes insulin-stimulated glucose transport (ISGT) and recruitment of GLUT4. *a*, G_{i2} -deficient adipocytes show markedly attenuated ISGT which cannot be explained solely by increased cAMP. *b*, Transgenic mice show a change in recruitment of GLUT4 to the plasma membrane in adipose tissue after insulin treatment *in vivo*.

METHODS. Adipocytes from control and transgenic mice were isolated by collagenase (1 mg ml^{-1}) digestion of epididymal fat pads in Krebs-Ringer HEPES (30 mM) buffer, pH 7.4, supplemented with BSA (3% w/v), adenosine deaminase (0.5 U ml^{-1}) and glucose (2.5 mM)²⁷. Normal and G_{i2} -deficient adipocytes ($2 \times 10^6 \text{ cells}$) were incubated in buffer with no additions (basal) or the indicated agents. After a 20-min incubation at



$37^\circ C$, [3H]-3-O-methyl-glucose uptake was measured essentially as described²⁷. Net uptake of [3H]-3-O-methyl glucose was calculated and initial uptake velocity determined as before²⁸. To assay GLUT4 recruitment, mice were treated with insulin and at the times indicated post-injection, their epididymal fat pads were removed and low-density microsomal membrane (L), high-density microsomal membrane (H), and plasma membrane fractions (P) were prepared²⁷. An aliquot of membrane protein ($25 \mu g$ protein per SDS-PAGE lane per gel) was subjected to SDS-polyacrylamide gel electrophoresis on a 10% polyacrylamide gel, and the separated proteins transferred to nitrocellulose. For immunoblot analysis, we used anti-GLUT4 antibodies (1:200 dilution; gift from H. Haspel)⁴. As a

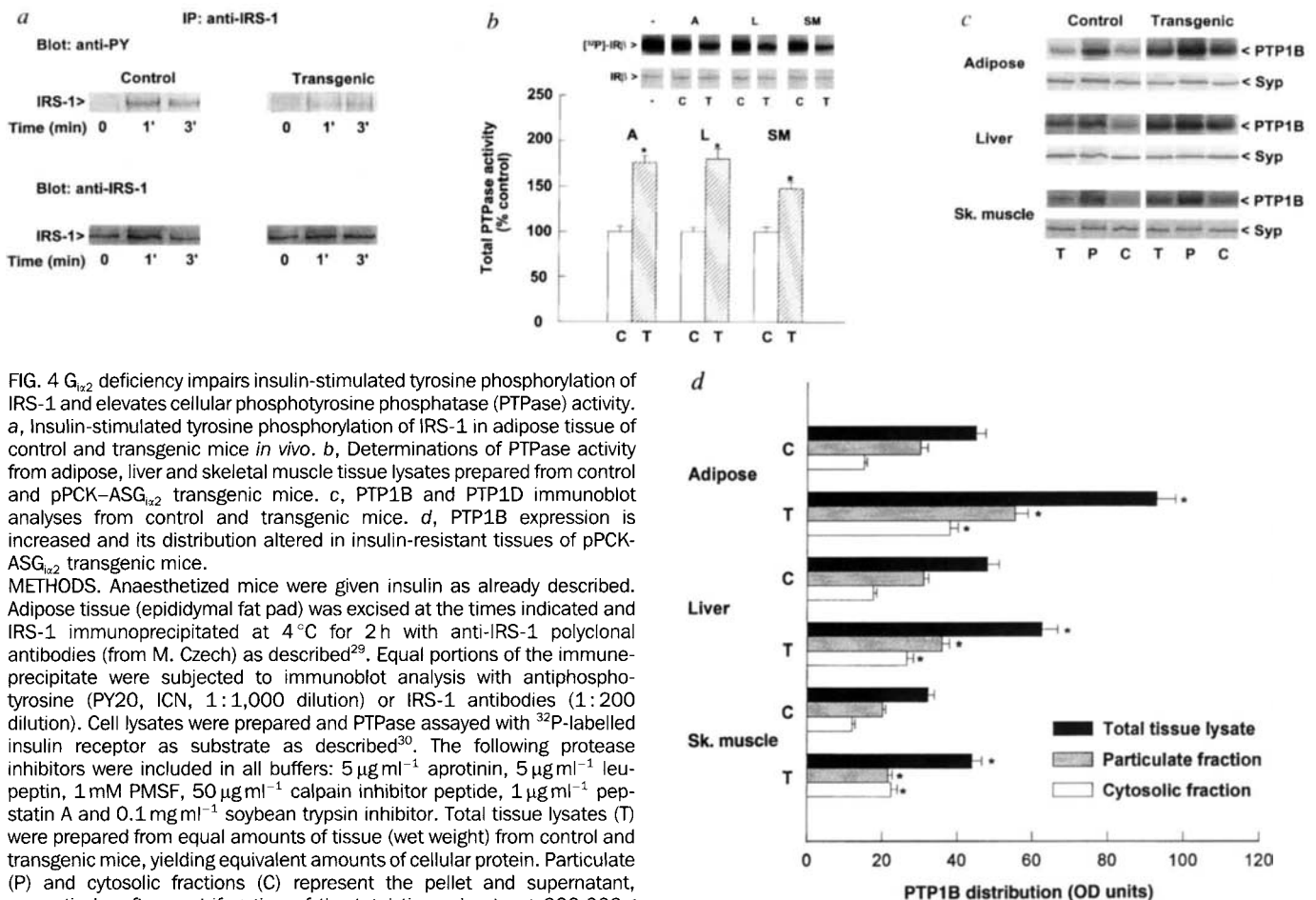


FIG. 4 G_{12} deficiency impairs insulin-stimulated tyrosine phosphorylation of IRS-1 and elevates cellular phosphotyrosine phosphatase (PTPase) activity. **a**, Insulin-stimulated tyrosine phosphorylation of IRS-1 in adipose tissue of control and transgenic mice *in vivo*. **b**, Determinations of PTPase activity from adipose, liver and skeletal muscle tissue lysates prepared from control and pPCK-ASG $_{12}$ transgenic mice. **c**, PTP1B and PTP1D immunoblot analyses from control and transgenic mice. **d**, PTP1B expression is increased and its distribution altered in insulin-resistant tissues of pPCK-ASG $_{12}$ transgenic mice.

METHODS. Anaesthetized mice were given insulin as already described. Adipose tissue (epididymal fat pad) was excised at the times indicated and IRS-1 immunoprecipitated at 4 °C for 2 h with anti-IRS-1 polyclonal antibodies (from M. Czech) as described²⁹. Equal portions of the immunoprecipitate were subjected to immunoblot analysis with antiphosphotyrosine (PY20, ICN, 1:1,000 dilution) or IRS-1 antibodies (1:200 dilution). Cell lysates were prepared and PTPase assayed with ³²P-labelled insulin receptor as substrate as described³⁰. The following protease inhibitors were included in all buffers: 5 μ g ml⁻¹ aprotinin, 5 μ g ml⁻¹ leupeptin, 1 mM PMSF, 50 μ g ml⁻¹ calpain inhibitor peptide, 1 μ g ml⁻¹ pepstatin A and 0.1 mg ml⁻¹ soybean trypsin inhibitor. Total tissue lysates (T) were prepared from equal amounts of tissue (wet weight) from control and transgenic mice, yielding equivalent amounts of cellular protein. Particulate (P) and cytosolic fractions (C) represent the pellet and supernatant, respectively, after centrifugation of the total tissue lysate at 200,000 g for 1 h at 4 °C. The amount of ³²P-labelled insulin added was equivalent for each sample (inset in **b**), and no significant proteolysis of the receptor was observed during the reactions. PTP1B and PTP1D immunoblot analyses (25 μ g protein per lane) were done as described⁴, using monoclonal antibodies specific for PTP1B and PTP1D (Transduction Lab). Total PTPase activity and distribution were calculated based upon determinations of protein in the total tissue lysate and subcellular fractions derived from same. Results are representative of at least three individual experiments performed on separate days with tissue from separate animals. The data in

b and **d** are mean values $\pm$ s.e.m. ($n = 3$); asterisks denote values significantly different from control ($P < 0.05$). The densitometric values (relative units of absorbance, means $\pm$ s.e.m., $n = 3$) for the anti-phosphotyrosine immunoblotting data in **a**, top, are as follows: control minus insulin, 0.04 ± 0.01 ; control + insulin at 1 min, 0.38 ± 0.02 ; and control + insulin at 3 min, 0.31 ± 0.01 . For pPCK-ASG $_{12}$ transgenic mice values were: transgenic minus insulin, 0.03 ± 0.01 ; transgenic + insulin at 1 min, 0.04 ± 0.01 ; and transgenic + insulin at 3 min, 0.04 ± 0.01 .

control, blots were probed with anti-insulin receptor antibodies (Transduction Labs) to ensure that equivalent amounts of membrane protein were loaded for each sample. GLUT4 immunoreactivity in plasma membrane fractions was quantified using a BioRad GS-670 imaging densitometer and Molecular Analyst software. In **a**, results are mean values $\pm$ s.e.m. ($n = 6$). Asterisks denote values significantly different from control ($P < 0.05$). Results presented in **b** are representative of three independent experiments, done on separate days with tissue from different animals. The densitometric values (relative units of absorbance, means $\pm$ s.e.m., $n = 3$) for the GLUT4 immunoblotting data from low-density microsomal (L), high-density microsomal (H), and plasma membrane-enriched (P) fractions are as follows. BDF1 control minus insulin: (L) 36.43 ± 3.20 , (H) 29.52 ± 2.12 , (P) 7.91 ± 0.65 ; BDF1 control plus insulin at 6 min: (L) 23.85 ± 1.37 , (H) 31.37 ± 0.82 , (P) 10.3 ± 0.47 ; BDF1 control plus insulin at 9 min: (L) 19.84 ± 0.93 , (H) 32.76 ± 1.70 , (P) 17.97 ± 1.19 ; and BDF1 control plus insulin at 12 min: (L) 9.97 ± 0.36 , (H) 30.49 ± 2.66 , (P) 28.34 ± 1.75 . For pPCK-ASG $_{12}$ transgenic mice the values are as follows. pPCK-ASG $_{12}$ transgenic minus insulin: (L) 26.08 ± 2.38 , (H) 29.34 ± 1.85 , (P) 11.92 ± 0.64 ; pPCK-ASG $_{12}$ transgenic plus insulin at 6 min: (L) 26.17 ± 1.27 , (H) 31.69 ± 3.42 , (P) 7.66 ± 0.69 ; pPCK-ASG $_{12}$ transgenic plus insulin at 9 min: (L) 29.87 ± 1.87 , (H) 28.06 ± 2.13 , (P) 10.92 ± 0.94 ; and pPCK-ASG $_{12}$ transgenic plus insulin at 12 min: (L) 32.66 ± 1.86 , (H) 27.90 ± 2.24 , (P) 8.28 ± 0.37 .

rodents¹²⁻¹⁵ and humans¹⁶, demonstrating increased PTPase activity and the release of PTP1B from its membrane-associated, endoplasmic reticular location¹⁷ to the cytosol. The mechanism of this redistribution is not known, although an increase in intracellular calcium¹⁸ or proteolytic cleavage of the C-terminal tail^{18,19} may be responsible. Other PTPases like PTP-LAR¹² may contribute to insulin resistance in G_{12} deficiency.

We have developed an antisense RNA expression system in which the specific role(s) of G proteins in several insulin-sensitive tissues can be explored *in vivo*. The runt, insulin-resistant phenotype of mice harbouring the pPCK-ASG $_{12}$ transgene is very similar to that reported for mice with a disrupted *IRS-1* gene^{20,21}. Our data are also consistent with reports of reduced levels of G_i in insulin-sensitive tissues of streptozotocin-treated, diabetic rats^{22,23}. G_{12} functions *in vivo* not only as a mediator of inhibitory adenylyl cyclase²⁻⁴, but also as a critical, positive regulator of insulin action, providing a link between the G-protein and tyrosine kinase signalling cascades. □

Received 14 September 1995; accepted 9 January 1996.

- Ullrich, A. & Schlessinger, J. *Cell* **61**, 203–212 (1990).
- Gilman, A. G. *Rev. Biochem.* **56**, 615–647 (1987).
- Moxham, C. M., Hod, Y. & Malbon, C. C. *Science* **260**, 991–995 (1993).

4. Moxham, C. M., Hod, Y. & Malbon, C. C. *Devl Gen.* **14**, 266–273 (1993).
5. White, M. F. & Kahn, C. R. *J. biol. Chem.* **269**, 1–4 (1994).
6. Reaven, G. M. *Physiol. Rev.* **75**, 473–486 (1995).
7. Cohen, P., Parker, P. J. & Woodgett, J. R. in *Molecular Basis of Insulin Action* (ed. Czech, M. P.) 213–233 (Plenum, New York, 1985).
8. Lavan, B. E. & Lienhard, G. E. *Biochem. Soc. Trans.* **22**, 676–680 (1994).
9. Cushman, S. W. & Wardzala, L. J. *J. biol. Chem.* **255**, 4758–4762 (1980).
10. Milarski, K. L. & Saltiel, A. R. *J. biol. Chem.* **269**, 21239–21243 (1994).
11. Noguchi, T. et al. *Molec. cell. Biol.* **14**, 6674–6682 (1994).
12. Ahmad, F. & Goldstein, B. *Am. J. Physiol.: Endocr. Metab.* **268**, E932–940 (1995).
13. Boylan, J. M. et al. *J. clin. Invest.* **90**, 174–179 (1992).
14. Begum, N., Sussman, K. E. & Draznin, B. *Diabetes* **40**, 1620–1629 (1991).
15. Meyerovitch, J., Backer, J. M. & Kahn, C. R. *J. clin. Invest.* **84**, 976–983 (1989).
16. McGuire, M. C. et al. *Diabetes* **40**, 939–942 (1991).
17. Mauro, L. J. et al. *Adv. Prot. Phosphatases* **7**, 393–411 (1993).
18. Frangioni, J. V. et al. *EMBO J.* **12**, 4843–4856 (1993).
19. Ezumi, Y., Takayama, H. & Okuma, M. *J. biol. Chem.* **270**, 11927–11934 (1995).
20. Tamemoto, H. et al. *Nature* **372**, 182–186 (1994).
21. Araki, E. et al. *Nature* **372**, 186–190 (1994).
22. Bushfield, M. et al. *Biochem. J.* **271**, 365–372 (1990).
23. Gawler, D., Milligan, G., Spiegel, A. M., Unson, C. G. & Houslay, M. D. *Nature* **327**, 224–232 (1987).
24. Wieland, O. *Methods of Enzymatic Analysis* (ed. O. Bergmeyer) Vol. 3, 1404–1414 (Academic, New York, 1974).
25. Guinovart, J. J. et al. *FEBS Lett.* **106**, 284–288 (1979).
26. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–269 (1951).
27. Karnieli, E. et al. *J. biol. Chem.* **256**, 4772–4777 (1981).
28. Foley, J. E., Cushman, S. W. & Salans, L. B. *Am. J. Phys.* **234**, E112–119 (1978).
29. Meyers, M. G. et al. *Molec. cell. Biol.* **14**, 3577–3587 (1994).
30. King, M. J. & Sale, G. J. *Biochem. J.* **256**, 893–902 (1988).

ACKNOWLEDGEMENTS. Supported by a training grant from the NIH (C.M.M.) and a grant from the American Cancer Society (C.C.M.).

DNA-binding properties of the yeast SWI/SNF complex

Janet Quinn^{*†}, Amy M. Fyrberg^{*},
Raymond W. Ganster[‡], Martin C. Schmidt[‡]
& Craig L. Peterson^{*§}

^{*}Program in Molecular Medicine and Department of Biochemistry and Molecular Biology, University of Massachusetts Medical Centre, Worcester, Massachusetts 01605, USA

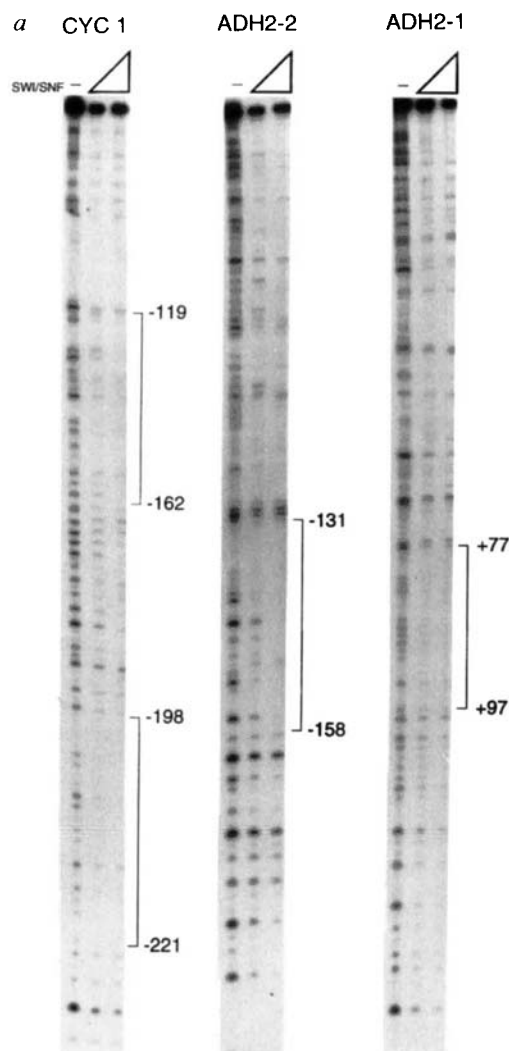
[‡]Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA

THE SWI/SNF complex is required for the enhancement of transcription by many transcriptional activators in yeast^{1,2}. Genetic and biochemical studies indicate that the complex facilitates activator function by antagonizing chromatin-mediated transcriptional repression^{3–6}. The absence of known DNA-binding motifs in several SWI/SNF subunits and the failure to identify SWI/SNF-dependent DNA-binding activities in crude yeast extracts have led to the belief that the complex does not bind DNA^{7,8}. Here we show that the SWI/SNF complex has a high affinity for DNA and that its DNA-binding properties are similar to those of proteins containing HMG-box domains⁹. The complex interacts with the minor groove of the DNA helix, binds synthetic four-way junction DNA, and introduces positive supercoils into relaxed plasmid DNA. These properties are likely to be important in the remodelling of chromatin structure by the SWI/SNF complex.

A gel retardation assay was used to test whether the purified yeast SWI/SNF complex binds DNA. The data in Fig. 1 show that the SWI/SNF complex binds with high affinity to promoter sequences from both SWI-dependent (*SUC2* and *ADH2*) and SWI-independent (*CYC1*) genes (Fig. 1a). Based on the concentration of SWI/SNF complex needed to bind half the probe DNA, we estimate the apparent binding constant to be $1–9 \times 10^{-9}$ M

(this value assumes 1–3 binding sites per probe molecule). We also found that the complex bound equally well to both promoter and coding-region sequences derived from the SWI-dependent gene, *HO*, demonstrating that the complex has no intrinsic affinity for regulatory sequences (results not shown). In all DNA-binding experiments, the addition of ATP had no effect on the formation of protein–DNA complexes (not shown). Although the complex does appear to be rather promiscuous for high-affinity DNA binding, it has some sequence specificity. For instance, all binding reactions contain a 1,000-fold molar excess of the double-stranded DNA poly(dG-dC)·poly(dG-dC). In addition, the complex binds poorly to the upstream activation sequence of the SWI-dependent *INO1* gene (Fig. 1a) and the 154-base-pair (bp) GAL4-binding sequence used in our previous study⁶ (data not shown). Furthermore, binding of the complex to *ADH2* and *CYC1* DNA probes resulted in protection of specific sequences from DNase (Fig. 2a). As in the gel retardation assays, addition of ATP to footprinting reactions had no effect (data not shown).

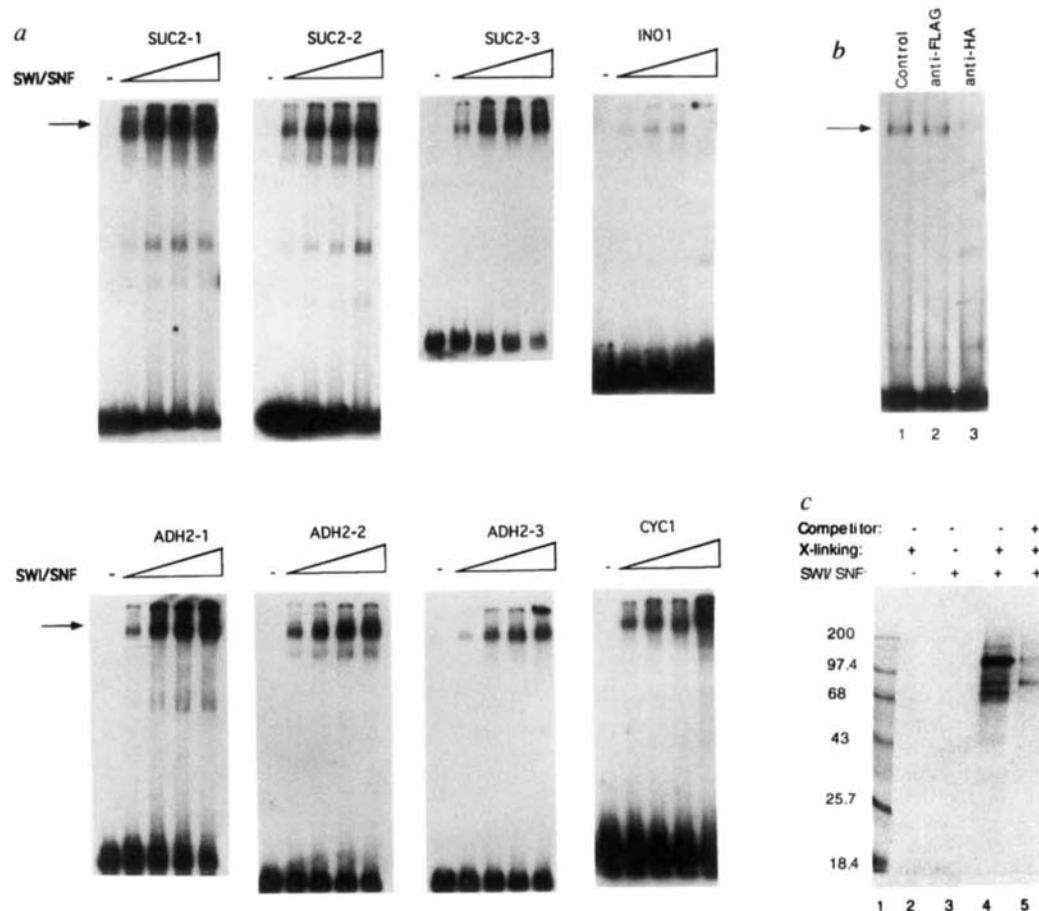
Four observations indicate that this DNA-binding activity is due to SWI/SNF complex. First, SWI/SNF complex can be fractionated on a native-DNA cellulose column (Fig. 1 legend). Second, DNA binding activity co-elutes with SWI/SNF complex from the final gel filtration column (data not shown). Third, fractions from a mock SWI/SNF purification do not contain detectable DNA-binding activity (data not shown). Fourth, addition of a monoclonal antibody directed against an engineered epitope at the C



[†] Present address: Departments of Biochemistry & Genetics and Medicine, University of Newcastle upon Tyne, NE2 4HH, UK.

[§] To whom correspondence should be addressed.

FIG. 1 The yeast SWI/SNF complex binds to DNA. **a**, Gel retardation analysis. Each panel shows a set of parallel gel retardation experiments in which increasing amounts of SWI/SNF complex (1.5 to 6.0 nM) were incubated with 0.1 nM 32 P-labelled promoter sequences derived from both SWI-dependent (*SUC2*, *ADH2*, *INO1*) and SWI-independent (*CYC1*) genes. Probes used in these experiments include: *SUC2*-1, *SUC2*-2 and *SUC2*-3, which contain bases -171 to +14, -171 to -386 and -386 to -712, respectively, of the *SUC2* promoter; *INO1* encompasses -126 to -333 of the *INO1* promoter; *ADH2*-1, *ADH2*-2 and *ADH2*-3 contain -54 to +116, -182 to -54, and -413 to -182, respectively, of the *ADH2* promoter; *CYC1* encompasses -249 to +11 of the *CYC1* promoter. **b**, Antibody supershifts. DNA-binding reactions were supplemented with 80 ng anti-FLAG monoclonal antibody (lane 2), 80 ng 12CA5 monoclonal antibody (lane 3) or received no antibody (lane 1). **c**, UV crosslinking of the SWI/SNF complex to the *SUC2*-3 probe. Lane 1, no UV; lane 2, no SWI/SNF; lane 3, +UV, +6 nM SWI/SNF; lane 4, identical to lane 3 but reactions contained a 20-fold molar excess of unlabelled probe. **METHODS.** DNA-binding assays (15 μ l) contained DNA-binding buffer (4 mM Tris-Cl, pH 7.4, 5 mM MgCl₂, 75 mM NaCl, 1 mM DTT, 0.05 mg ml⁻¹ BSA, 4% glycerol), 300 ng poly(dG-dC)·poly(dG-dC), 0.1 nM end-labelled 32 P-probe DNA, and 1.5 to 6 nM SWI/SNF complex. Protein-DNA complexes were resolved on 4% polyacrylamide (80:1 acrylamide:bisacrylamide ratio)/0.4 $\times$ TBE gels, and analysed by autoradiography. SWI/SNF complex was purified as described⁶, with an additional DNA-cellulose column step. Mock SWI/SNF purifications used strain CY340,



b
ADH2-1 +77 TTCCAAGCCAAAGCCCAACG⁺⁹⁷
ADH2-2 -158 AAATAGAGTGCAGTAGCGACTTTTTC⁻¹³¹
CYC1-a -162 ACGACACATGATCATATGGCATGCATGTGCTCTGTATGTATAT⁻¹¹⁹
CYC1-b -221 TAGCGTGGATGCCAGGCAACTT⁻¹⁹⁸

FIG. 2 a, Footprinting analysis of the yeast SWI/SNF complex. DNase I digestion patterns of the *CYC1*, *ADH2*-1 and *ADH2*-2 probes, either in the absence (lanes 1, 4 and 7) or in the presence of 6 nM (lanes 2, 5 and 8) or 12 nM (lanes 3, 6 and 9) SWI/SNF complex. **b**, Sequence comparison of SWI/SNF binding sites.

METHODS. Gel retardation assays were scaled up by a factor of two. Binding reactions were incubated for 30 min, followed by DNase I digestion (0.03 U) for 1 min at room temperature. Reactions were stopped by the addition of 5 volumes of STOP (5 mM Tris, pH 7.5, 175 mM NaCl, 10 mM EDTA, 3.5 M urea, 1% SDS), phenol/chloroform-extracted and ethanol-precipitated. Samples were analysed on 8% polyacrylamide/8M urea sequencing gels.

isogenic to strain CY396 (ref. 6), which does not express the SWI2-HA-6HIS fusion. Fractions from the Mono-Q step were diluted to 100 mM NaCl and loaded onto a 0.5 ml DNA-cellulose column pre-equilibrated in DNA-binding buffer. Bound protein was eluted with DNA-binding buffer containing 300 mM NaCl, and then subjected to gel filtration as described⁶. UV crosslinking: the *SUC2*-3 probe was internally labelled with [α - 32 P]dCTP, and bromodeoxyuridine was incorporated in place of dTTP to increase the efficiency of crosslinking. Standard DNA-binding reactions were UV-irradiated as described²⁵, treated with 1 unit DNase I for 30 min at 37 °C and analysed on 10% SDS-PAGE.

terminus of SWI2 disrupts and supershifts the DNA-protein complex (Fig. 1b, lane 3). Addition of a control antibody had no effect (Fig. 1b, lane 2). Ultraviolet crosslinking experiments were done to identify which of the 11 polypeptides^{6,10,11} of the SWI/SNF complex contact DNA. Three polypeptides were consistently crosslinked that corresponded to the 150K SWI1 polypeptide (*M_r* 150,000) and to two (p68 and p78) components of the complex that have not yet been cloned (Fig. 1c).

The DNA-binding properties of the SWI/SNF complex share many features of HMG-box domains. Like several HMG-box-containing polypeptides⁹, the SWI/SNF complex does not appear to recognize a consensus DNA sequence (Fig. 2b). The minor-groove-binding drugs distamycin A (Fig. 3a) and chromomycin A3 (data not shown), compete with SWI/SNF complex for DNA binding suggesting that, like HMG-box proteins¹²⁻¹⁴, the complex interacts (at least in part) with the minor groove. We also found that the SWI/SNF complex, like the HMG-box-domain-containing polypeptides HMG-1 (ref. 15) and UBF (ref. 16), exhibited a DNA-length dependence in binding assays (Fig. 3b).

One striking feature of HMG-box domains is their ability to recognize structured DNA⁹ such as cruciforms and four-way

junctions^{17,18}. As shown in Fig. 3c, d, the SWI/SNF complex binds with high affinity (apparent binding constant, 1 nM) to synthetic four-way-junction (4WJ) DNA but not to the duplex 'arms'. Binding to the synthetic 4WJ was also sensitive to low concentrations of distamycin (data not shown). The recognition of structured DNA may be crucial for SWI/SNF function: we found that 6 nM synthetic 4WJ was as effective as double-stranded plasmid DNA⁶ in stimulating the ATPase activity of the complex (> 6-fold), whereas the same concentration of the duplex 'arms' was completely ineffective. Interestingly, the structure formed in synthetic 4WJ DNA may mimic the crossover point where DNA enters and exits the nucleosome¹⁹.

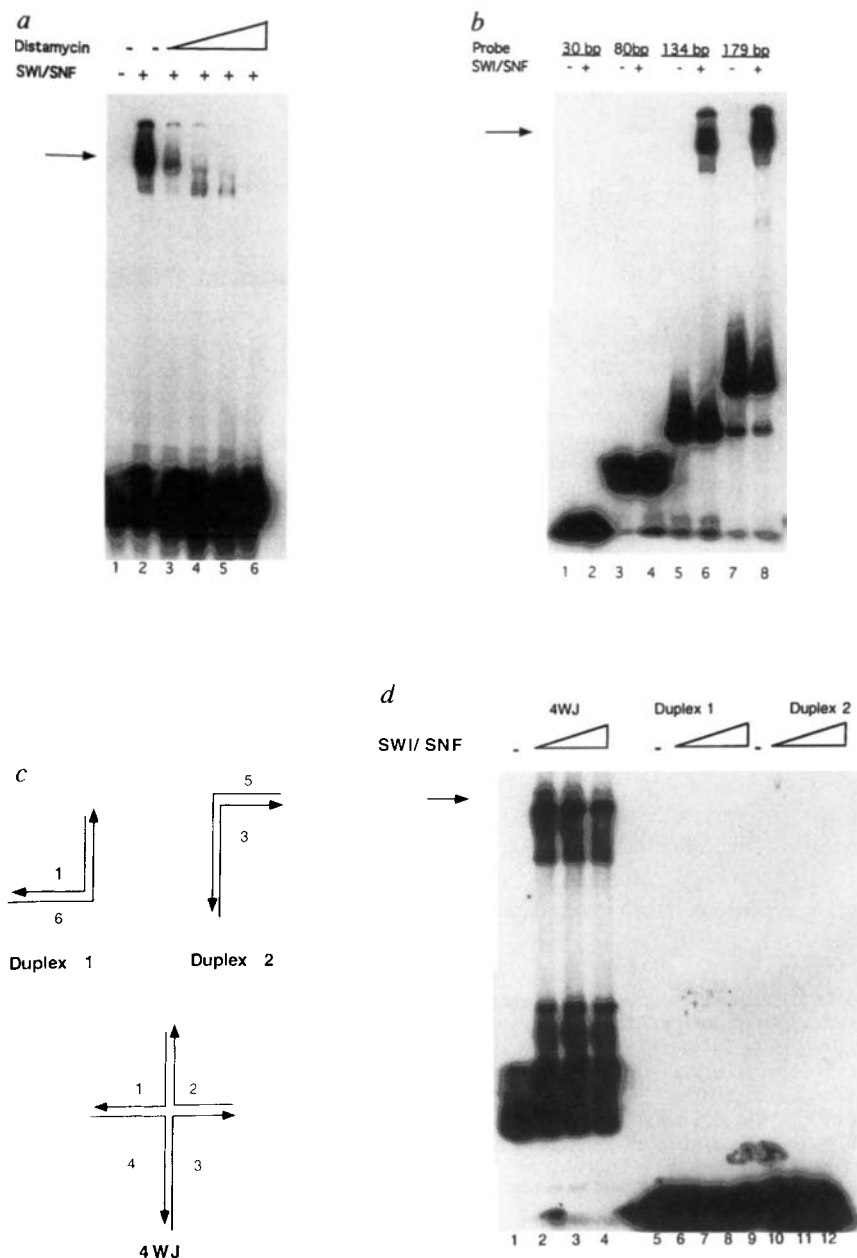
We also found that the SWI/SNF complex can introduce supercoils into circular DNA in the presence of bacterial topoisomerase I (Fig. 4). SWI/SNF complex is extremely active in this assay as it is able to drive a large change in linking number (> 10) even at a 1:1 ratio of SWI/SNF to plasmid (Fig. 4a, lanes 7 and 8). By comparison, the binding of about 200 molecules of HMG-1 per plasmid is required to introduce one superhelical turn²⁰. The SWI/

SNF complex did not introduce supercoils in the absence of topoisomerase I, nor did this activity require ATP. The observation that eukaryotic topoisomerases are able to remove SWI/SNF-induced supercoils (Fig. 4b; compare lanes 6 to 8 and 9) where *E. coli* topoisomerase I has no effect (Fig. 4b; compare lanes 6 and 7), indicates that the SWI/SNF complex introduces only positive supercoils in this assay. The electrophoretic mobility of the SWI/SNF-induced supercoils on a two-dimensional chloroquine gel is also consistent with positive supercoils (data not shown). In contrast, several HMG-box-containing polypeptides, including HMG-1 (ref. 20), mtTF1 (ref. 21) and ABF2 (ref. 22), induce negative supercoils into DNA in the presence of topoisomerase I.

What is the basis of the SWI/SNF complex-induced supercoiling? One possibility is that the SWI/SNF complex wraps the DNA template in a right-handed direction, introducing positive supercoils. But wrapping is unlikely because SWI/SNF binding does not induce the periodic hypersensitivity to DNase I that is typically observed for such DNA-protein complexes (Fig. 2). Alternatively, SWI/SNF may not wrap DNA but may use the

FIG. 3 The SWI/SNF complex has similar DNA-binding properties to HMG-box domains. **a**, The minor-groove-binding reagent distamycin A competes with the SWI/SNF complex for binding to DNA. Increasing amounts of distamycin A were added to DNA-binding reactions, which contained 0.1 nM labelled SUC2-3 fragment and 3 nM SWI/SNF complex, and its effect on binding assessed by gel retardation analysis. Lanes 1, probe DNA alone; 2, probe DNA and the SWI/SNF complex; 3–6, constant amount of SWI/SNF complex with increasing amounts of distamycin A (concentrations used were: lane 3, 0.1 μ M (molar ratio of reagent per bp, 0.4); lane 4, 1 μ M (molar ratio, 4); lane 5, 10 μ M (molar ratio, 40); lane 6, 100 μ M (molar ratio, 400)). **b**, Binding of SWI/SNF complex to DNA depends on the length of DNA. The complex (3 nM) was tested for binding to the –129 to –158 sequence of the *ADH2* promoter (Fig. 2b) which was contained on DNA fragments of different sizes. All DNA probes were present at 0.1 nM in the binding reactions. **c** and **d**, The SWI/SNF complex binds to synthetic cruciform DNA: **c**, the synthetic cruciform DNA (4WJ) and the control duplexes whose sequences are the same as each 'half' of the synthetic cruciform DNA. Numbers 1–6 denote oligonucleotides that were annealed to form the 4WJ (1–4) and control duplexes (1,6 and 5,3)²⁶. **d**, a gel retardation analysis of DNA-binding reactions in which the complex (lanes 2, 6, 10: 1.5 nM; lanes 3, 7, 11: 3 nM) was incubated with 0.1 nM cruciform DNA (lanes 2–4) or with 0.1 nM duplex molecules (lanes 6–8 and 10–12).

METHODS. To prepare probes for the length-dependence experiment, oligonucleotides containing bases –129 to –158 of the *ADH2* promoter were annealed to produce the 30-bp probe. The 80-bp probe was generated by PCR amplification of the –89 to –168 sequence of the *ADH2* promoter. Digestion of plasmid pHH26 with *Sall* or *EcoRI/HindIII* released fragments of 134 bp and 179 bp, respectively, which contained the –129 to –158 sequence.



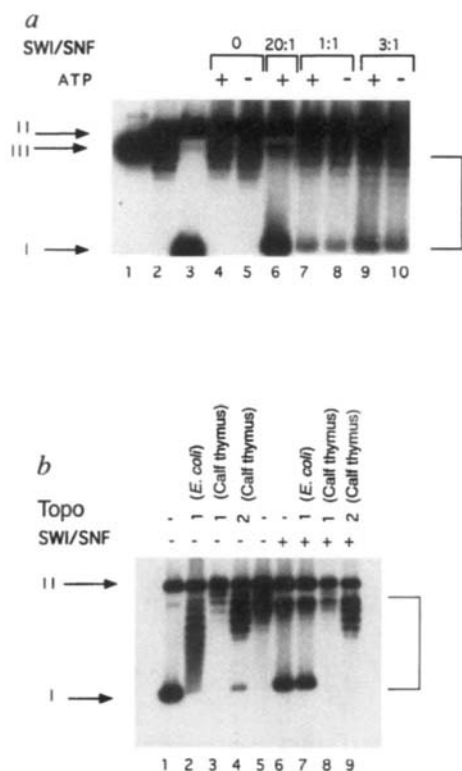


FIG. 4 SWI/SNF introduces positive supercoils into relaxed plasmid DNA in the presence of bacterial topoisomerase I. *a*, SWI/SNF induces supercoiling. Closed relaxed plasmid DNA (lane 2) was incubated with the SWI/SNF complex and bacterial topoisomerase I (topo I; 3 units). Molar ratios of SWI/SNF to plasmid are indicated above each set of lanes. A ratio of 20:1 corresponds to 3 nM SWI/SNF. Addition of ATP to a subset of the reactions is indicated. Arrows to the right denote topology standards, lane 1 contains linear plasmid (form III), lane 3 contains supercoiled plasmid (form I) and some nicked circles (comigrate with form II, closed relaxed DNA). The bracket to the right denotes SWI/SNF-induced topoisomers. *b*, SWI/SNF introduces positive supercoils. Plasmid DNA was supercoiled with SWI/SNF (3 nM) and bacterial topo I as in *a*. DNA was purified (lane 6) and retreated with either bacterial topo I (3 units; lane 7), calf thymus topo I (3 units; lane 8), or calf thymus topo II (7 units; lane 9). Lanes 1–4 show control reactions in which negatively supercoiled plasmid DNA (lane 1) was incubated with each topoisomerase under conditions identical to those for lanes 7–9. Lane 5 shows the starting relaxed substrate DNA. Reactions that contained calf thymus topo II contained 1 mM ATP.

METHODS. Supercoiling reactions (20 μ l) contained 1 $\times$ supercoiling buffer (20 mM HEPES, pH 7.5, 7 mM MgCl₂, 15 mM KCl, 0.5 mM DTT, 50 μ g per ml BSA), 50 ng pJH28, SWI/SNF, and topoisomerases where indicated. Reactions were incubated for 45 min at 30 °C, stopped with 80 μ l 1% SDS, 10 mM EDTA, 100 μ g ml⁻¹ proteinase K, 50 μ g ml⁻¹ tRNA, and incubated for 30 min at 37 °C. Samples were extracted with phenol/chloroform, ethanol-precipitated, and electrophoresed on 0.8% agarose gels without ethidium bromide and then Southern-blotted. DNA was purified and electrophoresed in the presence of chloroquine as described²⁷. Blots were probed with pJH28 labelled with [α -³²P]dCTP by random priming. Plasmid pJH28 contains SUC2 sequences from -1, 100 to +14 in plasmid pRS316.

ERRATA

Crystal structure of a G-protein $\beta\gamma$ dimer at 2.1 Å resolution

John Sondek, Andrew Bohm, David G. Lambright, Heidi E. Hamm & Paul B. Sigler

Nature 379, 369–374 (1996)

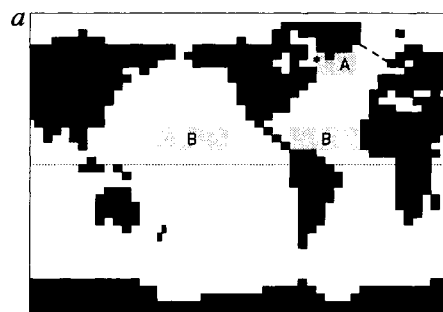
In this title, a typographical error caused the substitution of a subscripted 'A' for a hyphen in 'G-protein'. The correct title is given here. □

Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle

Stefan Rahmstorf

Nature 378, 145–149 (1995)

The shaded areas A and B of Fig. 1a of this Article were lost during printing. The correct figure is shown here. □



energy of DNA binding to change the helical twist, resulting in an overwinding of the DNA. Changes in helical twist may destabilize histone–DNA²³ as well as histone–histone interactions²⁴, so the ability of the SWI/SNF complex to modulate both DNA structure and topology may be important in SWI/SNF-dependent disruption of nucleosome structure. □

Received 20 September 1995; accepted 12 January 1996.

- Winston, F. & Carlson, M. *Trends Genet.* **8**, 387–391 (1992).
- Peterson, C. L. & Tamkun, J. W. *Trends biochem. Sci.* **20**, 143–146 (1995).
- Kruger, W. & Herskowitz, I. *Molec. cell. Biol.* **11**, 4135–4146 (1991).
- Hirschhorn, J. N., Brown, S. A., Clark, C. D. & Winston, F. *Genes Dev.* **6**, 228–2298 (1992).
- Kruger, W. *et al.* *Genes Dev.* **9**, 2770–2779 (1995).
- Côté, J., Quinn, J., Workman, J. L. & Peterson, C. L. *Science* **265**, 53–60 (1990).
- Laurent, B. C., Treitel, M. A. & Carlson, M. *Molec. cell. Biol.* **10**, 5616–5625 (1990).
- Peterson, C. L. & Herskowitz, I. *Cell* **68**, 573–583 (1992).
- Grosschedl, R., Giese, K. & Pagel, J. *Trends Genet.* **10**, 94–100 (1994).
- Cairns, B. R., Kim, Y.-J., Sayre, M. H., Laurent, B. C. & Kornberg, R. D. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1950–1954 (1994).
- Treich, I., Cairns, B. R., Santos, T., Brewster, E. & Carlson, M. *Molec. cell. Biol.* **15**, 4240–4248 (1995).
- Van de Wetering, M. & Clevers, H. *EMBO J.* **11**, 3039–3044 (1992).
- Copenhagen, G. P., Putnam, C. D., Denton, M. L. & Pikaard, C. S. *Nucleic Acids Res.* **22**, 2651–2657 (1994).
- Werner, M. H., Huth, J. R., Gronenborn, A. M. & Clore, G. M. *Cell* **81**, 705–714 (1995).
- Pil, P. M., Chow, C. S. & Lippard, S. J. *Proc. natn. Acad. Sci. U.S.A.* **90**, 9465–9469 (1993).
- Putnam, C. D. & Pikaard, C. S. *Molec. cell. Biol.* **12**, 4970–4980 (1992).
- Bianchi, M. E., Beltrame, M. & Paonessa, G. *Science* **243**, 1056–1059 (1989).
- Bruhn, S. L., Pil, P. M., Essigmann, J. M., Housman, D. E. & Lippard, S. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 2307–2311 (1992).
- Lilley, D. M. *J. Nature* **357**, 282–283 (1992).
- Stros, M., Stokrova, J. & Thomas, J. O. *Nucleic Acids Res.* **22**, 1044–1051 (1994).
- Fisher, R. P., Lisowsky, T., Parisi, M. A. & Clayton, D. A. *J. biol. Chem.* **267**, 3358–3367 (1992).
- Diffley, J. F. X. & Stillman, B. *J. biol. Chem.* **267**, 3368–3374 (1992).
- Lee, M.-S. & Garrard, W. T. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9675–9679 (1991).
- Clark, D. J. & Felsenfeld, G. *EMBO J.* **10**, 387–395 (1991).
- Peterson, C. L., Eaton, S. & Calame, K. *Molec. cell. Biol.* **8**, 4972–4980 (1988).
- Bianchi, M. E. *EMBO J.* **7**, 843–849 (1988).
- Knab, A. M., Fertala, J. & Bjornsti, M.-A. *J. biol. Chem.* **268**, 22322–22330 (1993).

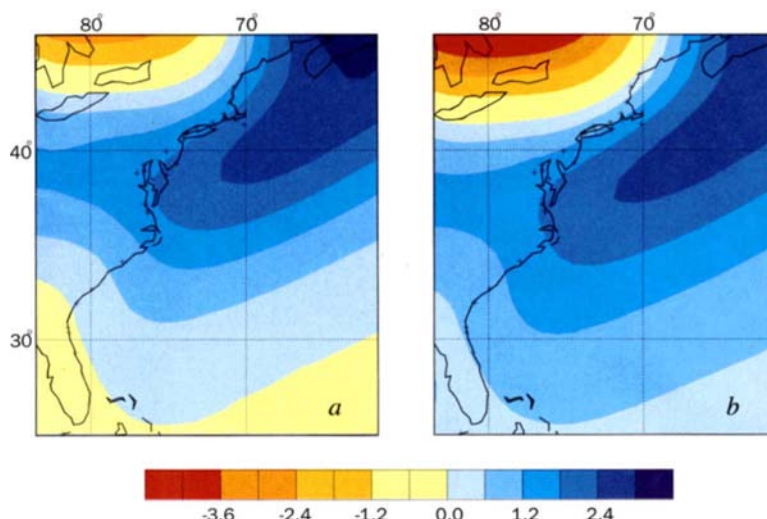
ACKNOWLEDGEMENTS. We thank E. T. Young for plasmid HH26; J. Hirschhorn and F. Winston for plasmid pJH28; J. Wang for *E. coli* topoisomerase I; G. Viglianti, S. Stern and members of C.L.P.s laboratory for discussion and for comments on the manuscript. This work was supported by grants to C.L.P. from the NIH and the March of Dimes Birth Defects Foundation. J.Q. is a Fellow and C.L.P. is a Scholar of the Leukemia Society of America.

Glacial isostatic adjustment and the anomalous tide gauge record of eastern North America

James L. Davis & Jerry X. Mitrovica

Nature 379, 331–333 (1996)

A LOW-RESOLUTION file for Fig. 3 was inadvertently used for publication. The figure is correctly reproduced here. □



Requirement of mammalian DNA polymerase- β in base-excision repair

Robert W. Sobol, Julie K. Horton, Ralf Kühn, Hua Gu, Rakesh K. Singhal, Rajendra Prasad, Klaus Rajewsky & Samuel H. Wilson

Nature 379, 183–186 (1996)

AN incorrect footnote symbol was used to indicate the name of the corresponding author (S.H.W.) of this Letter, which caused confusion over his proper affiliation. Correspondence should be addressed to S.H.W. at the Sealy Center for Molecular Science, University of Texas Medical Branch, Galveston, Texas 77555–1068, USA. □

CORRECTION

Dorsal cell fate specified by chick *Lmx1* during vertebrate limb development

Astrid Vogel, Concepción Rodríguez, Wayne Warnken & Juan Carlos Izpisua Belmonte

Nature 378, 716–720 (1995)

WE omitted to supply the Gene Accession number for chick *Lmx1*, which is U41823.

Also, expression of *C-Lmx1* was determined 24 hours, and not 48 hours as published, after removal of the apical ectodermal ridge or dorsal ectoderm. □

YOURS TO HAVE AND TO HOLD
BUT NOT TO COPY

The publication you are reading is protected by copyright law.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986